

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 39 Seconds
(without alignments)
1165.089 Million cell updates/sec

Title: US-09-401-636-4

Perfect score: 1858

Sequence: 1 EFRHHHHHTLSLPSPGVTI.....HEALPGSRTLEKSLHYSAGN 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1840	99.0	341	21	Immunogenic peptid
2	1595	85.8	341	21	Immunogenic peptid
3	1555	83.7	345	21	Immunogenic peptid
4	1533.5	83.6	342	21	Immunogenic peptid
5	1539	82.8	341	21	Immunogenic peptid
6	1526	82.1	341	21	Immunogenic peptid
7	1511.5	81.4	342	21	Immunogenic peptid
8	1051	56.6	340	21	Immunogenic peptid
9	1012	54.5	313	21	Rat IgE heavy chain
10	973	52.4	343	21	Rat immunoglobulin
					Platyus IgE heavy

11	940.5	50.6	421	23	AAU80300	Mouse IgE heavy ch
12	938.5	50.5	426	17	AAU97753	Canine IgE. Canis
13	936	50.4	332	23	AAU80297	Mouse IgE heavy ch
14	936	50.4	332	23	AAU80298	Murine IgE heavy c
15	936	50.4	332	23	AAU80299	Murine IgE heavy c
16	934.5	50.3	417	18	AAW23067	Canine IgE heavy c
17	929.5	50.0	432	23	AAU80294	Murine IgE heavy c
18	925	49.8	343	23	AAU80295	Murine IgE heavy c
19	925	49.8	343	23	AAU80296	Murine IgE heavy c
20	922.5	49.7	312	21	AAV79995	Dog immunoglobulin
21	898	48.3	424	23	AAU50103	Equine IgE heavy c
22	896	48.2	424	23	AAU50104	Equine IgE heavy c
23	862.5	46.4	561	20	AAV17415	Mouse immunoglobul
24	847	45.6	313	21	AAV79997	Mouse immunoglobul
25	745	40.1	323	23	AAU80284	Human IgE heavy ch
26	745	40.1	323	23	AAU80285	Human IgE C2-C3-C4
27	745	40.1	323	23	AAU80286	Human IgE C2-C3-C4
28	745	40.1	324	16	AAU83559	Fc(epsilon) CH2'-C
29	745	40.1	325	16	AAU75225	Human IgE Fc chain
30	745	40.1	325	16	AAU77241	Human IgE heavy ch
31	745	40.1	331	21	AAU803642	Human IgE heavy ch
32	745	40.1	367	9	AAU80291	Interleukin-2/IgE
33	745	40.1	428	23	AAU80283	Human IgE heavy ch
34	745	40.1	428	23	AAU50940	Human IgE epsilon
35	745	40.1	428	23	AAU47863	Human Ig-E heavy c
36	743	40.0	330	23	AAU80289	Human IgE C2-C3-C4
37	742	39.9	325	21	AAV79994	Human immunoglobul
38	741	39.9	325	16	AAU83582	CH2 to CH4 of huma
39	741	39.9	336	23	AAU80288	Human IgE heavy ch
40	741	39.9	441	23	AAU80287	Human IgE heavy ch
41	737	39.7	493	5	AAU40065	Sequence of human
42	726	39.1	315	16	AAU85582	Fc(epsilon) CH2'-C
43	709	38.2	428	14	AAU42950	Human IgE heavy ch
44	525.5	28.3	245	14	AAU38858	Anti-allergic chim
45	520.5	28.0	468	22	AAU69120	Canine IgG heavy c

ALIGNMENTS

RESULT 1

AAU806206
ID AAB06206 standard; protein; 341 AA.

XX AAB06206;

XX AAB06206;

XX 22-NOV-2000 (first entry)

XX Immunogenic peptide consisting of opossum CH2, rat CH3 and opossum CH4.

XX Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;

XX asthma; eczema; immunogenic peptide.

XX Chimeric - Didelphis virginiana.

XX Chimeric - Rattus sp.

XX WO200025722-A2.

XX 11-MAY-2000.

XX 21-OCT-1999; 99WO-SE01896.

XX 02-NOV-1998; 98US-0106652.

XX 22-SEP-1999; 99US-0401636.

XX (RESI-) RESISTENTIA PHARM AB.

XX Hellman LT;

XX WPI; 2000-365342/31.

XX Immunogenic polypeptides useful for preventing the harmful effects of

XX immunoglobulin E in mammals -

XX Disclosure; Fig 2; 50pp; English.

CC The present sequence is an immunogenic peptide consisting of the heavy chain constant regions 2 and 4 of the opossum IgE and the heavy chain constant region 3 from the rat. It was shown to cause a stronger polyclonal anti-self IgE response than peptides consisting of the same regions from one mammal. Immunogenic peptides, particularly those consisting of different heavy chain constant regions, can be used for vaccination in humans, against bacterial and viral infections and allergies, such as asthma, fur, pollen and food allergies and eczema.

XX Sequence 341 AA;

Query Match 99.0%; Score 1840; DB 21; Length 341;
Best Local Similarity 99.4%; Pred. No. 1.8e-144;
Matches 339; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60

Qy 61 LVDGQEAENLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTVCHVKNGSIFEDSSR 120
Db 61 LVDGQEAENLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTVCHVKNGSIFEDSSR 120

Qy 121 RCDDEPRGVITYLIPSPDLIYENGTPKLTCLVLDLSEENITVTVRERKKSIGASQ 180
Db 121 RCDDEPRGVITYLIPSPDLIYENGTPKLTCLVLDLSEENITVTVRERKKSIGASQ 180

Qy 181 RSTKHHTATTSITSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRLAPEVYM 240
Db 181 RSTKHHTATTSITSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRLAPEVYM 240

Qy 241 LPPSPETGTRTVTCLIRGFYPSISVQWLPNNEEDTGHHTTRPKQDGTDPSPFLY 300
Db 241 LPPSPETGTRTVTCLIRGFYPSISVQWLPNNEEDTGHHTTRPKQDGTDPSPFLY 300

Qy 301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 2
AAB06208
ID AAB06208 standard; protein; 341 AA.

XX AC AAB06208;
XX DT 22-NOV-2000 (first entry)
XX DE Immunogenic peptide consisting of opossum CH2, dog CH3 and opossum CH4.
XX Dog; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
XX asthma; eczema; immunogenic peptide.
XX OS Chimeric - Didelphis virginiana.
XX OS Chimeric - Canis sp.
XX FN WO200025722-A2.
XX PD 11-MAY-2000.
XX PF 21-OCT-1999; 99WO-SE01896.
XX PR 02-NOV-1998; 98US-0106652.
XX PR 22-SEP-1999; 99US-0401636.
XX PA (RESI-) RESISTENTIA PHARM AB.
XX PI .Hellman LT;
XX DR . WPI; 2000-365342/31.

XX Immunogenic polypeptides useful for preventing the harmful effects of immunoglobulin E in mammals -
PT Disclosure; Fig 2; 50pp; English.

CC The present sequence is an immunogenic peptide consisting of the heavy chain constant regions 2 and 4 of the opossum IgE and the heavy chain constant region 3 from the dog. It was shown to cause a stronger polyclonal anti-self IgE response than peptides consisting of the same regions from one mammal. Immunogenic peptides, particularly those consisting of different heavy chain constant regions, can be used for vaccination in humans, against bacterial and viral infections and allergies, such as asthma, fur, pollen and food allergies and eczema.

XX Sequence 341 AA;

Query Match 85.8%; Score 1595; DB 21; Length 341;
Best Local Similarity 85.0%; Pred. No. 3.8e-124;
Matches 290; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60

Qy 61 LVDGQEAENLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTVCHVKNGSIFEDSSR 120
Db 61 LVDGQEAENLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTVCHVKNGSIFEDSSR 120

Qy 121 RCDDEPRGVITYLIPSPDLIYENGTPKLTCLVLDLSEENITVTVRERKKSIGASQ 180
Db 121 RCDDEPRGVITYLIPSPDLIYENGTPKLTCLVLDLSEENITVTVRERKKSIGASQ 180

Qy 181 RSTKHHTATTSITSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRLAPEVYM 240
Db 181 NKXDFHNGTITVSTLTPVNTDMEGETYYCRVTHPLPKDIVERIAKLPKRLAPEVYM 240

Qy 241 LPPSPETGTRTVTCLIRGFYPSISVQWLPNNEEDTGHHTTRPKQDGTDPSPFLY 300
Db 241 LPPSPETGTRTVTCLIRGFYPSISVQWLPNNEEDTGHHTTRPKQDGTDPSPFLY 300

Qy 301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 3
AAB06207
ID AAB06207 standard; protein; 345 AA.

XX AC AAB06207;
XX DT 22-NOV-2000 (first entry)
XX DE Immunogenic peptide consisting of opossum CH2, pig CH3 and opossum CH4.
XX Pig; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
XX asthma; eczema; immunogenic peptide.
XX OS Chimeric - Didelphis virginiana.
XX OS Chimeric - Sus scrofa;
XX FN WO200025722-A2.
XX PD 11-MAY-2000.
XX PF 21-OCT-1999; 99WO-SE01896.
XX PR 02-NOV-1998; 98US-0106652.
XX PR 22-SEP-1999; 99US-0401636.
XX PA (RESI-) RESISTENTIA PHARM AB.
XX PI .Hellman LT;
XX DR . WPI; 2000-365342/31.

PI Hellman LT;
 XX WPI; 2000-365342/31.
 XX Immunogenic polypeptides useful for preventing the harmful effects of
 XX immunoglobulin E in mammals -
 XX Disclosure; Fig 2; 50pp; English.
 XX The present sequence is an immunogenic peptide consisting of the heavy
 XX chain constant regions 2 and 4 of the opossum IgE and the heavy chain
 XX constant region 3 from the pig. It was shown to cause a stronger
 XX polyclonal anti-self IgE response than peptides consisting of the same
 XX regions from one mammal. Immunogenic peptides, particularly those
 XX consisting of different heavy chain constant regions, can be used for
 XX vaccination in humans, against bacterial and viral infections and
 XX allergies, such as asthma, fur, pollen and food allergies and eczema.
 XX Sequence 345 AA;
 SQ Query Match 83.7%; Score 1555; DB 21; Length 345;
 Best Local Similarity 84.1%; Pred. No. 8e-121;
 Matches 290; Conservative 13; Mismatches 38; Indels 4; Gaps 2;
 QY 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTV 60
 DB 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTV 60
 QY 61 LVDQEAENLFPYTRPKREGGQTFSLQSEVNIQGMSSNTYTCHVKHNGSIFEDSSR 120
 DB 61 LVDQEAENLFPYTRPKREGGQTFSLQSEVNIQGMSSNTYTCHVKHNGSIFEDSSR 120
 QY 121 RC-SDEPRGVITYLIPSPDLVYENGTPKLTCLVLDLSEENITVTVWRERKKS1--G 176
 DB 121 RCTAESERGVAVLSPTFDLVLVHKSFKLTCLVLDLASENVNLLMSRENKGVILPP 180
 QY 177 SASQRTKHHATTSTISILPVDADKWIEGEGYOCRVDPHPKPIVRSITKLPKRLAP 236
 DB 181 PGPPVVKPQNGTFSATSTLPVNSDWIEGTYGCVNTHPDLKPKILRSIKLPGKRLAP 240
 QY 237 EYVMLPSPETGTTRVTCLIRGFYPSSEISVQWLPNNEEDHTGHHTTRPKDHGTDPS 296
 DB 241 EYVMLPSPETGTTRVTCLIRGFYPSSEISVQWLPNNEEDHTGHHTTRPKDHGTDPS 300
 QY 297 FFLLSRMLVNKSIWEKGNLVCRCVVHEALPGSRTLEKSLHYSAGN 341
 DB 301 FFLLSRMLVNKSIWEKGNLVCRCVVHEALPGSRTLEKSLHYSAGN 345

RESULT 4
 AAB06205
 ID AAB06205 standard; protein; 342 AA.
 XX AAB06205;
 AC AAB06205;
 XX 22-NOV-2000 (first entry)
 DT 22-NOV-2000 (first entry)
 DE Immunogenic peptide consisting of opossum CH2, human CH3 and opossum CH4.
 XX Human; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
 KW asthma; eczema; immunogenic peptide.
 XX Chimeric - Didelphis virginiana.
 OS Chimeric - Homo sapiens.
 XX WO200025722-A2.
 PN 11-MAY-2000.
 XX 21-OCT-1999; 99WO-SE01896.
 XX 02-NOV-1998; 98US-0106652.
 PR 22-SEP-1999; 99US-0401636.

XX (RESI-) RESISTENTIA PHARM AB.
 XX PA Hellman LT;
 PI WPI; 2000-365342/31.
 XX Immunogenic polypeptides useful for preventing the harmful effects of
 XX immunoglobulin E in mammals -
 XX Disclosure; Fig 2; 50pp; English.
 XX The present sequence is an immunogenic peptide consisting of the heavy
 XX chain constant regions 2 and 4 of the opossum IgE and the heavy chain
 XX constant region 3 from the human. It was shown to cause a stronger
 XX polyclonal anti-self IgE response than peptides consisting of the same
 XX regions from one mammal. Immunogenic peptides, particularly those
 XX consisting of different heavy chain constant regions, can be used for
 XX vaccination in humans, against bacterial and viral infections and
 XX allergies, such as asthma, fur, pollen and food allergies and eczema.
 XX Sequence 342 AA;
 SQ Query Match 83.6%; Score 1553.5; DB 21; Length 342;
 Best Local Similarity 82.7%; Pred. No. 1.1e-120;
 Matches 283; Conservative 21; Mismatches 37; Indels 1; Gaps 1;
 QY 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTV 60
 DB 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTV 60
 QY 61 LVDQEAENLFPYTRPKREGGQTFSLQSEVNIQGMSSNTYTCHVKHNGSIFEDSSR 120
 DB 61 LVDQEAENLFPYTRPKREGGQTFSLQSEVNIQGMSSNTYTCHVKHNGSIFEDSSR 120
 QY 121 RCDSDPRGVITYLIPSPDLVYENGTPKLTCLVLDLSEENITVTVWRERKKSISGAS 179
 DB 121 KCADSNRPGVAVLSRPSFDLFRKSPFTITCLVVDLAPSKGTNLTWASRSGKPVNHST 180
 QY 180 QRSTKHHATTSTISILPVDADKWIEGEGYOCRVDPHPKPIVRSITKLPKRLAPEVY 239
 DB 181 RKEKQKNGTLVTSTLPGVTRDIEGETQCRVTHPHLPRALMRSTTKLPKRLAPEVY 240
 QY 240 MLPSPETGTTRVTCLIRGFYPSSEISVQWLPNNEEDHTGHHTTRPKDHGTDPSFFL 299
 DB 241 MLPSPETGTTRVTCLIRGFYPSSEISVQWLPNNEEDHTGHHTTRPKDHGTDPSFFL 300
 QY 300 YSRMLVNKSIWEKGNLVCRCVVHEALPGSRTLEKSLHYSAGN 341
 DB 301 YSRMLVNKSIWEKGNLVCRCVVHEALPGSRTLEKSLHYSAGN 342

RESULT 5
 AAB06202
 ID AAB06202 standard; protein; 341 AA.
 XX AAB06202;
 AC AAB06202;
 XX 22-NOV-2000 (first entry)
 DT 22-NOV-2000 (first entry)
 DE Immunogenic peptide consisting of opossum CH2, mouse CH3 and opossum CH4.
 XX Mouse; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
 KW asthma; eczema; immunogenic peptide.
 XX Chimeric - Didelphis virginiana.
 OS Chimeric - Mus sp.
 XX Key Location/Qualifiers
 FT Misc-difference 1..341
 FT /label= OTHER
 FT /note= "Xaa=unknown"

PN WO200025722-A2.
 XX 11-MAY-2000.
 XX 21-OCT-1999; 99WO-SE01896.
 XX 02-NOV-1999; 98US-0106652.
 XX 22-SEP-1999; 99US-0401636.
 XX (RESI-) RESISTENTIA PHARM AB.
 XX Hellman LT;
 PI WPI; 2000-365342/31.
 DR Immunogenic polypeptides useful for preventing the harmful effects of
 PT immunoglobulin E in mammals -
 XX Disclosure; Fig 2; 50pp; English.
 PS The present sequence is an immunogenic peptide consisting of the heavy
 CC chain constant regions 2 and 4 of the opossum IGE and the heavy chain
 CC constant region 3 from the mouse. It was shown to cause a stronger
 CC polyclonal anti-self IGE response than peptides consisting of the same
 CC regions from one mammal. Immunogenic peptides, particularly those
 CC consisting of different heavy chain constant regions, can be used for
 CC vaccination in humans, against bacterial and viral infections and
 CC allergies, such as asthma, fur, pollen and food allergies and eczema.
 XX Sequence 341 AA;
 SQ Query Match 82.8%; Score 1539; DB 21; Length 341;
 Best Local Similarity 85.9%; Pred. No. 1.7e-119;
 Matches 293; Conservative 12; Mismatches 34; Indels 2; Gaps 2;
 QY 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHTW 60
 Db 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHTW 60
 QY 61 LVDSQEAENLPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
 Db 61 LVDSQEAENLPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
 QY 121 RCDDEPRGVITYLPPSPDLQYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQ 180
 Db 121 RCDDEPRGVITYLPPSPDLQYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQ 180
 QY 181 RSTKHHTATTSILPVDADKMTIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVYM 239
 Db 181 RSTKHHTATTSILPVDADKMTIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVYM 239
 QY 240 MLPSPSEETGTRTITVCLIRGFYFSEISVQMLPNNDEEDHTGHTTTRPKDGHGTDPSFPL 299
 Db 240 MLPSPSEETGTRTITVCLIRGFYFSEISVQMLPNNDEEDHTGHTTTRPKDGHGTDPSFPL 299
 QY 300 YSRMLVNSIWEKGNLVTCTRVVHEALPGSRITLKSLSHYSAG 340
 Db 300 YSRMLVNSIWEKGNLVTCTRVVHEALPGSRITLKSLSHYSAG 340
 RESULT 6
 AAB03644
 ID AAB03644 standard; protein; 341 AA.
 XX AAB03644;
 AC AAB03644;
 XX 22-NOV-2000 (first entry)
 DT Opossum IGE heavy chain constant regions 2, 3 and 4.
 DE Opossum IGE heavy chain constant regions 2, 3 and 4.
 KW Opossum; immunoglobulin E; IGE; vaccination; infection; allergy;
 XX asthma; eczema; immunogenic peptide.

OS Didelphis virginiana.
 XX WO200025722-A2.
 XX 11-MAY-2000.
 XX 21-OCT-1999; 99WO-SE01896.
 XX 02-NOV-1999; 98US-0106652.
 XX 22-SEP-1999; 99US-0401636.
 XX (RESI-) RESISTENTIA PHARM AB.
 XX Hellman LT;
 PI WPI; 2000-365342/31.
 DR Immunogenic polypeptides useful for preventing the harmful effects of
 PT immunoglobulin E in mammals -
 XX Disclosure; Fig 1; 50pp; English.
 PS The present sequence is an immunogenic peptide consisting of the heavy
 CC chain constant regions 2, 3 and 4 of the opossum IGE. It was used to
 CC construct a number of immunogenic peptides which consisted of regions of
 CC IGE from different mammals, which appear to cause a stronger polyclonal
 CC anti-self IGE response than peptides consisting of the same regions from
 CC one mammal. Immunogenic peptides, particularly those consisting of
 CC different heavy chain constant regions, can be used for vaccination
 CC in humans, against bacterial and viral infections and allergies, such
 CC as asthma, fur, pollen and food allergies and eczema.
 XX Sequence 341 AA;
 SQ Query Match 82.1%; Score 1526; DB 21; Length 341;
 Best Local Similarity 80.4%; Pred. No. 2e-118;
 Matches 274; Conservative 29; Mismatches 38; Indels 0; Gaps 0;
 QY 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHTW 60
 Db 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHTW 60
 QY 61 LVDSQEAENLPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
 Db 61 LVDSQEAENLPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
 QY 121 RCDDEPRGVITYLPPSPDLQYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQ 180
 Db 121 RCDDEPRGVITYLPPSPDLQYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQ 180
 QY 181 RSTKHHTATTSILPVDADKMTIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVYM 240
 Db 181 RSTKHHTATTSILPVDADKMTIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVYM 240
 QY 241 MLPSPSEETGTRTITVCLIRGFYFSEISVQMLPNNDEEDHTGHTTTRPKDGHGTDPSFPL 300
 Db 241 MLPSPSEETGTRTITVCLIRGFYFSEISVQMLPNNDEEDHTGHTTTRPKDGHGTDPSFPL 300
 QY 301 YSRMLVNSIWEKGNLVTCTRVVHEALPGSRITLKSLSHYSAG 341
 Db 301 YSRMLVNSIWEKGNLVTCTRVVHEALPGSRITLKSLSHYSAG 341
 RESULT 7
 AAB06201
 ID AAB06201 standard; protein; 342 AA.
 XX AAB06201;
 AC AAB06201;
 XX 22-NOV-2000 (first entry)
 DT Immunogenic peptide consisting of opossum CH2, CH4 and rat/opossum CH3.
 DE Immunogenic peptide consisting of opossum CH2, CH4 and rat/opossum CH3.
 XX

KW Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
 KW asthma; eczema; immunogenic peptide.
 XX
 OS Chimeric - Didelphis virginiana.
 OS Chimeric - Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..342
 FT /label= OTHER
 FT /note= "Xaa=unknown"
 XX
 XX WO200025722-A2.
 XX
 XX 11-MAY-2000.
 XX
 XX 21-OCT-1999; 99WO-SE01896.
 XX
 XX 02-NOV-1998; 98US-0106652.
 PR 22-SEP-1999; 99US-0401636.
 XX
 XX (RESI-) RESISTENTIA PHARM AB.
 XX
 XX Hellman LT;
 XX
 XX WPI; 2000-365342/31.
 XX
 XX Immunogenic polypeptides useful for preventing the harmful effects of
 XX immunoglobulin E in mammals -
 XX
 XX Disclosure; Fig 2; 50pp; English.
 XX
 CC The present sequence is an immunogenic peptide consisting of the heavy
 CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
 CC constant region 3 created from a combination of the one from the rat
 CC and the one from the opossum. It was shown to cause a stronger
 CC polyclonal anti-self IgE response than peptides consisting of the same
 CC regions from one mammal. Immunogenic peptides, particularly those
 CC consisting of different heavy chain constant regions, can be used for
 CC vaccination in humans, against bacterial and viral infections and
 CC allergies, such as asthma, fur, pollen and food allergies and eczema.
 XX
 XX Sequence 342 AA;
 SQ
 Query Match 81.4%; Score 1511.5; DB 21; Length 342;
 Best Local Similarity 83.6%; Pred. No. 3.2e-117;
 Matches 285; Conservative 16; Mismatches 39; Indels 1; Gaps 1;
 QY 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHTW 60
 DB 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHTW 60
 QY 61 LVDGQAEALPPTTRPKREGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
 DB 61 LVDGQAEALPPTTRPKREGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
 QY 121 RCDSDPRGVITLIPPSFLDYENGTPKLTCLVLDLSEENITVWVRERKKSIGSASQ 180
 DB 121 RCDSDPRGVITLIPPSFLDYENGTPKLTCLVLDLSEENITVWVRERKKSIGSASQ 180
 QY 181 RSTKH-HAATSIITSLPVDKADWIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPVY 239
 DB 181 LWKEXYHGTFXTSHLPVETDDWTEGXTYTXLLESDPMIVILTIPTISALPGKRLAPVY 240
 QY 240 MLPSPETGTRVTCILIRGFYPSISVQWLPNNEDHTGHHTTTPQKHGTDPPSFL 299
 DB 241 MLPSPETGTRVTCILIRGFYPSISVQWLPNNEDHTGHHTTTPQKHGTDPPSFL 300
 QY 300 YSRMLVHKSIEWKGHLNLTCTRVVHEALPGSRTLEKSLHYSAG 340
 DB 301 YSRMLVHKSIEWKGHLNLTCTRVVHEALPGSRTLEKSLHYSAG 341

RESULT 8

AAB03643
 ID AAB03643 standard; protein; 340 AA.
 XX
 AC AAB03643;
 XX
 DT 22-NOV-2000 (first entry)
 XX
 DE Rat IgE heavy chain constant regions 2, 3 and 4.
 XX
 KW Rat; immunoglobulin E; IgE; vaccination; infection; allergy;
 KW asthma; eczema; immunogenic peptide.
 XX
 OS Rattus sp.
 OS
 PN WO200025722-A2.
 XX
 XX 11-MAY-2000.
 XX
 XX 21-OCT-1999; 99WO-SE01896.
 PF
 XX 02-NOV-1998; 98US-0106652.
 PR 22-SEP-1999; 99US-0401636.
 XX
 XX (RESI-) RESISTENTIA PHARM AB.
 XX
 XX Hellman LT;
 XX
 XX WPI; 2000-365342/31.
 XX
 XX Immunogenic polypeptides useful for preventing the harmful effects of
 XX immunoglobulin E in mammals -
 XX
 XX Disclosure; Fig 1; 50pp; English.
 XX
 CC The present sequence is an immunogenic peptide consisting of the
 CC heavy chain constant regions 2, 3 and 4 of the rat IgE. It was used to
 CC construct a number of immunogenic peptides which consisted of regions of
 CC IgE from different mammals, which appear to cause a stronger polyclonal
 CC anti-self IgE response than peptides consisting of the same regions from
 CC one mammal. Immunogenic peptides, particularly those consisting of
 CC different heavy chain constant regions, can be used for vaccination
 CC in humans, against bacterial and viral infections and allergies, such
 CC as asthma, fur, pollen and food allergies and eczema.
 XX
 XX Sequence 340 AA;
 SQ
 Query Match 56.6%; Score 1051; DB 21; Length 340;
 Best Local Similarity 62.6%; Pred. No. 5.2e-79;
 Matches 206; Conservative 35; Mismatches 78; Indels 10; Gaps 3;
 QY 17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHTVWLVDGQAEALFPYVTR 76
 DB 8 PVTITKTVDLLHSSCDPNA-FHSTIQLYCFVYGHQNDVSIHWMDDRKI-----YETH 61
 QY 77 PK----REGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSRRCSDDEPRGVIT 132
 DB 62 AQNVLIKEEGKLASTYSLNITQQWMSSTFTCKVTSGENYWAHTRRCSDDEPRGVIT 121
 QY 133 YLIPPSFLDYENGTPKLTCLVLDLSEENITVWVRERKKSIGSASQSRSTKHHHTSI 192
 DB 122 YLIPPSFLDYENGTPKLTCLVLDLSEENITVWVRERKKSIGSASQSRSTKHHHTSI 181
 QY 193 TSILPVDKADWIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPVYMLPSPETGTR 252
 DB 182 TSILPVDKADWIEGEGYQCRVDHPFPKPIVRSITKAPGRSAFVYVFLPPEEEKDKR 241
 QY 253 TVTCLIRGFYPSISVQWLPNNEDHTGHHTTTPQKHGTDPPSFLYSRMLVHKSIEWKG 312
 DB 242 TLTCLIQNFPPEDISVQWLQDSKLI PKSQHSTTTPKLYNGSNQRFIFSRLEVTKALWTQ 301
 QY 313 GNLVTCRVVHEALPGSRTLEKSLHYSAG 341
 DB 302 TKQFTCRVHIEALREPRKLERTISKSLGN 330

```
RESULT 9
ID AAY79996 standard; Protein; 313 AA.
XX
AC AAY79996;
XX
DT 15-MAY-2000 (first entry)
XX
DE Rat immunoglobulin E epsilon heavy chain SEQ ID NO:3.
XX
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Rattus sp.
XX
PN WO9967293-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13959.
XX
PR 20-JUN-1998; 98US-0100287.
XX
PA (UNBII-) UNITED BIOMEDICAL INC.
XX
PI Wang CY, Walfield AM;
XX
DR WPI; 2000-160578/14.
XX
PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy
XX
PS Example 1; Page 66-68; 155pp; English.
XX
CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX
SQ Sequence 313 AA;

Query Match 54.5%; Score 1012; DB 21; Length 313;
Best Local Similarity 63.3%; Pred. No. 8e-76; Mismatches 10; Gaps 3;
Matches 198; Conservative 32; Indels 73;

QY 17 PVTIIPPTVKLPHSSCDPRGDAHSTIQLCLVSGFSPAKVHVWMLVDGQEAENLFFVYTR 76
DB 3 PWNITKPTVLLHSSCDPNA-FHSTIQLYCFYGHQNDVSIHMLMDRKI-----YDTH 56

QY 77 PK-----REGQTFISQSEVNITQGMMSNTYTCVHKVNGSIFEDSSRRCSDDPRGVIT 132
DB 57 AQNVLIKEEGKLASTYSRLNITQQQMGSEFTCKVTQGENYWAHTRRCSDDPRGVIT 116

QY 133 YLIPSPDLVYENGTPKLTCLVLDSESENIYTVWRERKKSIGSASORSTKHHHTTSI 192
DB 117 YLIPSPDLVYENGTPKLTCLVLDSESENIYTVWRERKKSIGSASORSTKHHHTTSI 176

QY 193 TSILPVDKDWIEGEGYQCRVDHPHFKPIVRISITKLPKRLAPEVVMPLPSPETGTR 252
DB 177 TSILPVDKDWIEGEGYQCRVDHPHFKPIVRISITKALGLRSAFEVVVFLPPEEKNKR 236
QY 253 TVTCLIRGFYPSBISVQWLNNBEDHTGHHTTRPQKDHGTDPSPFFLYSRLMNVKNSIWEK 312
DB 237 TLTCLIQNFPEDISVQWLQDSKLIPKQSHSTTTPLTKNGSNQRFFFISRLVTKALWTQ 296
QY 313 GNLVTCRVVHEAL 325
DB 297 TKQFTCRVIHEAL 309

RESULT 10
AAB06204
ID AAB06204 standard; protein; 343 AA.
XX
AC AAB06204;
XX
DT 22-NOV-2000. (first entry)
XX
DE Platypus IgE heavy chain constant regions 2, 3 and 4.
XX
KW Platypus; immunoglobulin E; IgE; vaccination; infection; allergy;
KW asthma; eczema; immunogenic peptide.
XX
OS Ornithorhynchus anatinus.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..343
FT /label= OTHER
FT /note= "Xaa=unknown"
XX
PN WO200025722-A2.
XX
PD 11-MAY-2000.
XX
PF 21-OCT-1999; 99WO-SE01896.
XX
PR 02-NOV-1998; 98US-0106652.
XX
PR 22-SEP-1999; 99US-0401636.
XX
PA (RESI-) RESISTENTIA PHARM AB.
XX
PI Hellman LT;
XX
DR WPI; 2000-365342/31.
XX
PT Immunogenic polypeptides useful for preventing the harmful effects of
PT immunoglobulin E in mammals -
XX
PS Disclosure; Fig 2; 50pp; English.
XX
CC The present sequence is an immunogenic peptide consisting of the heavy
CC chain constant regions 2, 3 and 4 of the platypus IgE. It was used to
CC construct a number of immunogenic peptides which consisted of regions of
CC IgE from different mammals, which appear to cause a stronger polyclonal
CC anti-self IgE response than peptides consisting of the same regions from
CC one mammal. Immunogenic peptides, particularly those consisting of
CC different heavy chain constant regions, can be used for vaccination
CC in humans, against bacterial and viral infections and allergies, such
CC as asthma, fur, pollen and food allergies and eczema.
XX
SQ Sequence 343 AA;

Query Match 52.4%; Score 973; DB 21; Length 343;
Best Local Similarity 52.3%; Pred. No. 1.6e-72;
Matches 180; Conservative 58; Mismatches 100; Indels 6; Gaps 4;

QY 1 EFHHHHHTLSESGPVTIIPPTVKLPHSSCDPRGDAHSTIQLCLVSGFSPAKVHVW 60
DB 1 EFHHHHHTTEVYSDSSK-DPIPTVKLHSSCDPRGDSQASIELLCLITGYSPAGIQVDW 59
```


CC viral systems. The IgE can be used in drug development (e.g.
CC small molecule screening, assay development and anti-IgE
CC antibody generation). Fragments of IgE can be used in vaccines
CC or to prevent IgE-mediated hypersensitivity. The new sequence
CC information permits targeted modulation of IgE-mediated immune
CC responses.

XX SQ Sequence 426 AA;

Query Match 50.5%; Score 938.5; DB 17; Length 426;
Best Local Similarity 53.2%; Pred. No. 1.5e-69;
Matches 173; Conservative 53; Mismatches 96; Indels 3; Gaps 2;

QY 18 VTIIPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVTVLVDGQAEENLFPYTTRP 77
DB 102 LNFIPPTVKLFHSSCNFVGDTHITLQLCLISGYVPGMEVWLVDGQAKNIFPYTAPG 161
QY 78 KREGQOTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSRRCSDDPRGVITYLIPP 137
DB 162 TKEGNVT-STHSELNITQGEVWSQXTYTCQVYQGFTEKDEARKCESDPRGVITLSP 220
QY 138 SPLDIYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQSRSTKHHHATTSITSLP 197
DB 221 SPLDIYVHKAPKITCLVLDLATMEGMNLTWVRESKEPVNPGLNKDKHFNGTITVTSLP 280
QY 198 VDADKWIEGEGYQCVDPHPPKPIVRSITKLPKGLAPEVYMLPPSPPEETGTTT -TVT 255
DB 281 VNTNDWIEGETYCYRVTHPLPKDIVRSIAKAPGRAPDVYFLPPEEQGQKDRVTIT 340
QY 256 CLIRGFYFSEISVQWLPNNEEDHTGHTTTRPQKHGTDPSFFLYSRMLVNKSIWEKGNL 315
DB 341 CLIQNFFPADISVQWLNDSPITQDTTGTGPHKVSGRPAFFIFSRLEVSRLVDWEQKNK 400
QY 316 VTCRVVHEALPGSRITLKSLSHYSAG 340
DB 401 FTCQVVIHEALSGSRILQKWSKTPG 425

RESULT 13

AAU80297
ID AAU80297 standard; Protein; 332 AA.

XX AC AAU80297;

XX DT 30-JUL-2002 (first entry)

XX DE Mouse IgE heavy chain C2-C3-C4 domain with MIGIS fragment.
XX IGE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;
XX antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IGE;
XX vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
XX heavy chain C domain; MIGIS.

XX OS Mus sp.

XX OS Synthetic.

XX PN WO200220038-A2.

XX XX 14-MAR-2002.

XX PF 06-SEP-2001; 2001WO-DK00579.

XX PR 06-SEP-2000; 2000DK-0001326.

XX PR 15-SEP-2000; 2000US-232831P.

XX XX (PHAR-) PHARMEXA AS.

XX PI Klynsner S, Von Hoegen P, Voldborg B, Gautam A;

XX DR WPI; 2002-383033/41.

XX PT *Inducing immune response against autologous immunoglobulin E in an
XX animal, by effecting simultaneous presentation of cytotoxic T

PT lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT -
XX Examples; Page 137-138; 151pp; English.

XX This invention relates to a novel method for inducing an immune response
CC against autologous immunoglobulin E (IGE) in an animal. The method
CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC (CTL) epitope and/or B-cell epitope derived from IGE, and T helper cell
CC (TH) epitope which is foreign to the animal, by antigen
CC presenting cells (APCs) of the animal's immune system. The epitopes
CC of the invention may be used as a vaccine against allergic diseases. The
CC method of the invention is useful for inducing an immune response
CC against autologous IGE in an animal. This method is useful for downregulating
CC autologous IGE in the animal. This method is useful for the prevention
CC and treatment of allergic diseases such as anaphylaxis, allergic
CC rhinitis, asthma and atopic dermatitis. The present sequence represents
CC the mouse IGE heavy chain C2-C3-C4 domain with the MIGIS fragment used
CC to create the epitopes used in the method of the invention.

SQ Sequence 332 AA;

Query Match 50.4%; Score 936; DB 23; Length 332;
Best Local Similarity 56.7%; Pred. No. 1.8e-69;
Matches 185; Conservative 46; Mismatches 91; Indels 4; Gaps 4;

QY 17 PVTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVTVLVDGQAEENLFPYTTR 76
DB 4 PVNITEPTLELHSSCDPNA-FHSTIQLCYFIYGHILNDVSVLMDREITDTLAQTVL 62
QY 77 PKREGGOTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSRRCSDDPRGVITYLIP 136
DB 63 IKEE-GKLASTCSKLNITEQQWMSEFTCKVTSQGVVDYLAHTRCPDHPERGVITYLIP 121
QY 137 PSLDLYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQSRSTKH-HATTSITSI 195
DB 122 PSLDLYQNGAPKLTCLVLDLSEKYNVTWNOEKTSV-SASOWYTKHNNATTSITSI 180
QY 196 LPVDAKWIEGEGYQCVDPHPPKPIVRSITKLPKGLAPEVYMLPPSPPEETGTTT 255
DB 181 LPVDAKWIEGEGYQCVDPHPPKPIVRSITKLPKGLAPEVYMLPPSPPEETGTTT 240
QY 256 CLIRGFYFSEISVQWLPNNEEDHTGHTTTRPQKHGTDPSFFLYSRMLVNKSIWEKGNL 315
DB 241 CLIQNFFPADISVQWLNDSPITQDTTGTGPHKVSGRPAFFIFSRLEVAKTLWTQSK 300
QY 316 VTCRVVHEALPGSRITLKSLSHYSAG 341
DB 301 FTCQVVIHEALQKPRKLEKTTISLGN 326

RESULT 14

AAU80298
ID AAU80298 standard; Protein; 332 AA.

XX AC AAU80298;

XX DT 30-JUL-2002 (first entry)

XX DE Murine IgE heavy chain C2-C3-C4 for mammalian expression.
XX IGE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;
XX antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IGE;
XX vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
XX heavy chain C domain.

XX OS Mus sp.

XX OS Synthetic.

XX PN WO200220038-A2.

XX XX 14-MAR-2002.

XX PD

PF 06-SEP-2001; 2001WO-DK00579.
XX
XX 06-SEP-2000; 2000DK-0001326.
PR 15-SEP-2000; 2000US-232831P.
XX
XX (PHAR-) PHARMEXA AS.
PA
PI Klyner S, Von Hoegen P, Voldborg B, Gautam A;
XX
XX WPI; 2002-383033/41.
DR N-PSDB; ABK51141.
XX
XX Inducing immune response against autologous immunoglobulin E in an
PT animal, by effecting simultaneous presentation of cytotoxic T
PT lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT
XX
XX Disclosure; Page 140-141; 151pp; English.
PS
XX This invention relates to a novel method for inducing an immune response
CC against autologous immunoglobulin E (IgE) in an animal. The method
CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC epitopes (TH epitope) which is foreign to the animal, by antigen
CC presenting cells (APCs) of the animal's immune system. The epitopes
CC of the invention may be used as a vaccine against allergic diseases. The
CC method of the invention is useful for inducing an immune response
CC against autologous IgE in an animal, which is useful for downregulating
CC autologous IgE in the animal. This method is useful in the prevention
CC and treatment of allergic diseases such as anaphylaxis, allergic
CC rhinitis, asthma and atopic dermatitis. The present sequence represents
CC the mouse IgE heavy chain C2-C3-C4 domain optimised for a mammalian
CC expression system used to create the epitopes used in the method of the
CC invention.
XX
XX Sequence 332 AA;
SQ
Query Match 50.4%; Score 936; DB 23; Length 332;
Best Local Similarity 56.7%; Pred. No. 1.8e-69;
Matches 185; Conservative 46; Mismatches 91; Indels 4; Gaps 4;
QY 17 PVTIPTTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVTVLVDGQAEENLFPVYTR 76
DB 4 PVNITEPTLELLHSSCDPNA-FHSTIQLCYFIYGHILNDVSVWLMDDREITDLAQTVL 62
QY 77 PKREGGQTFSLQSEVNITQGWMSNTYTCHVKGNGSIFEDSSRRCSDEPRGVITYLIP 136
DB 63 IKER-GKLASTCSKLNITEQWMSESTFTCKVTSQGVLYLAHTRCPDHPHPRGVITYLIP 121
QY 137 PSLDLYENGTPKLTCLVLDLESEENITVTVRERKKSIGSASQSTKH-HATTSTISI 195
DB 122 PSLDLYQNGAPKLTCLVLDLESEKNVNTWNOEKTSV-SASQWYTKHNNATTSTISI 180
QY 196 LPVDAKDWEIGYQCRVDHPFPKPIVRSITKLPGLAPVYMLPSPSETGTRTVT 255
DB 181 LPVAKDWIEGYQCIVDHPDPKPIVRSITKTPGQSAPEVYVFPPEESEDKRTLT 240
QY 256 CLIRGFYSEISVQWLPNNEEDHTGHTTTPQKDHGTDPSFFLYSRMLVNSKIWEKGNL 315
DB 241 CLIQNFPEDSVQWLGKGLINSQHSHTTTPKNSNGNSQGFIFSRLEVAKTLWTQKQ 300
QY 316 VTCRVVHEALPGSRTLKSLHYSAGN 341
DB 301 FTCQVHEALQKPKLENTISTISLGN 326
RESULT 15
AAU80299
ID AAU80299 standard; Protein; 332 AA.
XX
XX
AC AAU80299;
XX
DT 30-JUL-2002 (first entry)

XX
DE Murine IgE heavy chain C2-C3-C4 for E.Coli expression.
XX
KW IgE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;
KW antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;
KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;
KW heavy chain C domain.
OS
OS Mus sp.
XX Synthetic.
XX
PN WO200220038-A2.
XX
PD 14-MAR-2002.
XX
XX 06-SEP-2001; 2001WO-DK00579.
PF
XX 06-SEP-2000; 2000DK-0001326.
PR 15-SEP-2000; 2000US-232831P.
XX
XX (PHAR-) PHARMEXA AS.
XX
PI Klyner S, Von Hoegen P, Voldborg B, Gautam A;
XX
XX WPI; 2002-383033/41.
DR N-PSDB; ABK51142.
XX
XX Inducing immune response against autologous immunoglobulin E in an
PT animal, by effecting simultaneous presentation of cytotoxic T
PT lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin
PT
XX
XX Disclosure; Page 144-145; 151pp; English.
PS
XX This invention relates to a novel method for inducing an immune response
CC against autologous immunoglobulin E (IgE) in an animal. The method
CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte
CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell
CC epitopes (TH epitope) which is foreign to the animal, by antigen
CC presenting cells (APCs) of the animal's immune system. The epitopes
CC of the invention may be used as a vaccine against allergic diseases. The
CC method of the invention is useful for inducing an immune response
CC against autologous IgE in an animal, which is useful for downregulating
CC autologous IgE in the animal. This method is useful in the prevention
CC and treatment of allergic diseases such as anaphylaxis, allergic
CC rhinitis, asthma and atopic dermatitis. The present sequence represents
CC the mouse IgE heavy chain C2-C3-C4 domain optimised for an E.Coli
CC expression system used to create the epitopes used in the method of the
CC invention.
XX
XX Sequence 332 AA;
SQ
Query Match 50.4%; Score 936; DB 23; Length 332;
Best Local Similarity 56.7%; Pred. No. 1.8e-69;
Matches 185; Conservative 46; Mismatches 91; Indels 4; Gaps 4;
QY 17 PVTIPTTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVTVLVDGQAEENLFPVYTR 76
DB 4 PVNITEPTLELLHSSCDPNA-FHSTIQLCYFIYGHILNDVSVWLMDDREITDLAQTVL 62
QY 77 PKREGGQTFSLQSEVNITQGWMSNTYTCHVKGNGSIFEDSSRRCSDEPRGVITYLIP 136
DB 63 IKER-GKLASTCSKLNITEQWMSESTFTCKVTSQGVLYLAHTRCPDHPHPRGVITYLIP 121
QY 137 PSLDLYENGTPKLTCLVLDLESEENITVTVRERKKSIGSASQSTKH-HATTSTISI 195
DB 122 PSLDLYQNGAPKLTCLVLDLESEKNVNTWNOEKTSV-SASQWYTKHNNATTSTISI 180
QY 196 LPVDAKDWEIGYQCRVDHPFPKPIVRSITKLPGLAPVYMLPSPSETGTRTVT 255
DB 181 LPVAKDWIEGYQCIVDHPDPKPIVRSITKTPGQSAPEVYVFPPEESEDKRTLT 240
QY 256 CLIRGFYSEISVQWLPNNEEDHTGHTTTPQKDHGTDPSFFLYSRMLVNSKIWEKGNL 315

Db 241 CLIQFFPDISVQWLGDKLISNSQHSHTTTPKSNQSGFFIFSRLEVAKTLWTQKQ 300
Qy 316 VTCRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 FTCQVIHEALQPKRKLEKTISTSLGN 326

Search completed: July 9, 2003, 13:41:48
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:40:31 ; Search time 53 Seconds
(without alignments)
749.040 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHTLSLPESGPVTI.....HEALPGSRTLKSLHYSAGN 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pap.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pap.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB.pap.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pap.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1858	100.0	341	9	US-10-176-664-4
2	1858	100.0	341	10	US-09-401-636-4
3	1840	99.0	341	9	US-10-176-664-9
4	1840	99.0	341	10	US-09-401-636-9
5	1720	92.6	341	9	US-10-176-664-6
6	1720	92.6	341	10	US-09-401-636-6
7	1691.5	91.0	342	9	US-10-176-664-5
8	1691.5	91.0	342	10	US-09-401-636-5
9	1595	85.8	341	9	US-10-176-664-11
10	1595	85.8	341	10	US-09-401-636-11
11	1555	83.7	345	9	US-10-176-664-10
12	1555	83.7	345	10	US-09-401-636-10
13	1553.5	83.6	342	9	US-10-176-664-8
14	1553.5	83.6	342	10	US-09-401-636-8
15	1529	82.3	341	9	US-10-176-664-3
16	1529	82.3	341	10	US-09-401-636-3
17	1465	78.8	446	9	US-10-214-524-32
18	1254	67.5	427	9	US-10-214-524-36
19	1051	56.6	340	9	US-10-176-664-2

20	1051	56.6	340	10	US-09-401-636-2	Sequence 2, Appli
21	1051	56.6	428	9	US-10-214-524-34	Sequence 34, Appli
22	1025	55.2	343	9	US-10-176-664-7	Sequence 7, Appli
23	1025	55.2	343	10	US-09-401-636-7	Sequence 7, Appli
24	1020.5	54.9	346	9	US-10-152-190-14	Sequence 14, Appli
25	1018.5	54.8	577	9	US-10-214-524-29	Sequence 29, Appli
26	950.5	51.2	426	9	US-10-214-524-28	Sequence 28, Appli
27	940.5	50.6	421	9	US-09-499-375A-28	Sequence 28, Appli
28	940.5	50.6	421	9	US-10-214-524-31	Sequence 31, Appli
29	937.5	50.5	496	9	US-10-214-524-25	Sequence 25, Appli
30	936	50.4	332	9	US-09-499-375A-23	Sequence 23, Appli
31	936	50.4	332	9	US-09-499-375A-25	Sequence 25, Appli
32	936	50.4	332	9	US-09-499-375A-27	Sequence 27, Appli
33	931.5	50.1	431	9	US-09-479-614-14	Sequence 14, Appli
34	931.5	50.1	496	9	US-09-479-614-2	Sequence 2, Appli
35	931.5	50.1	496	9	US-09-479-614-29	Sequence 29, Appli
36	929.5	50.0	432	9	US-09-499-375A-19	Sequence 19, Appli
37	925	49.8	343	9	US-09-499-375A-20	Sequence 20, Appli
38	925	49.8	343	9	US-09-499-375A-22	Sequence 22, Appli
39	865.5	46.6	569	9	US-10-214-524-30	Sequence 30, Appli
40	853	45.9	567	9	US-10-214-524-33	Sequence 33, Appli
41	835	44.9	563	9	US-10-214-524-35	Sequence 35, Appli
42	785	42.2	426	9	US-10-214-524-27	Sequence 27, Appli
43	774.5	41.7	346	9	US-10-152-190-10	Sequence 10, Appli
44	759	40.9	347	9	US-10-152-190-12	Sequence 12, Appli
45	752	40.5	569	9	US-09-847-208-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-176-664-4
; Sequence 4, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-4

Query Match	100.0%	Score 1858;	DB 9;	Length 341;
Best Local Similarity	100.0%	Pred. No. 4.9e-122;		
Matches	341;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	EFHHHHHTLSLPESGPVTIIPVTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW	60	
Db	1	EFHHHHHTLSLPESGPVTIIPVTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW	60	
Qy	61	LVDSQEAENLPYTRTRKREGGQTFSLQSEVNITQGGWMSNTYTCHVKHNGSIFEDSSR	120	
Db	61	LVDSQEAENLPYTRTRKREGGQTFSLQSEVNITQGGWMSNTYTCHVKHNGSIFEDSSR	120	
Qy	121	RCSDDPRGVITYLIPSPDLVYENGTPKLTCLVLDSESENIITVTVWRKKKIGSASQ	180	
Db	121	RCSDDPRGVITYLIPSPDLVYENGTPKLTCLVLDSESENIITVTVWRKKKIGSASQ	180	
Qy	181	RSTKHHTATTSITSLPVDKADKWTEGSGYQCRVDHPFPKPIVRSITKLPGKRLAPEVYM	240	
Db	181	RSTKHHTATTSITSLPVDKADKWTEGSGYQCRVDHPFPKPIVRSITKLPGKRLAPEVYM	240	

Db 181 RSTKHHHTTSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYM 240
Qy 241 LPPSPETGTRTTCILRGYPSSISVQWLPNNEEDHTGHTTTRPKDHGTDPSPFLY 300
Db 241 LPPSPETGTRTTCILRGYPSSISVQWLPNNEEDHTGHTTTRPKDHGTDPSPFLY 300
Qy 301 SRMLVNSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 SRMLVNSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 2

US-09-401-636-4
; Sequence 4, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-4

Query Match 100.0%; Score 1858; DB 10; Length 341;
Best Local Similarity 100.0%; Pred. No. 4.9e-122;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Qy 61 LVDQEAENLFPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKNGSIFEDSSR 120
Db 61 LVDQEAENLFPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKNGSIFEDSSR 120
Qy 121 RCDDEPRGVITLPPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKKSIGASQ 180
Db 121 RCDDEPRGVITLPPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKKSIGASQ 180
Qy 181 RSTKHHHTTSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYM 240
Db 181 RSTKHHHTTSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYM 240
Qy 241 LPPSPETGTRTTCILRGYPSSISVQWLPNNEEDHTGHTTTRPKDHGTDPSPFLY 300
Db 241 LPPSPETGTRTTCILRGYPSSISVQWLPNNEEDHTGHTTTRPKDHGTDPSPFLY 300
Qy 301 SRMLVNSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 SRMLVNSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 3

US-10-176-664-9
; Sequence 9, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636

; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-9

Query Match 99.0%; Score 1840; DB 9; Length 341;
Best Local Similarity 99.4%; Pred. No. 8.8e-121;
Matches 339; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Qy 61 LVDQEAENLFPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKNGSIFEDSSR 120
Db 61 LVDQEAENLFPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKNGSIFEDSSR 120
Qy 121 RCDDEPRGVITLPPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKKSIGASQ 180
Db 121 RCDDEPRGVITLPPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKKSIGASQ 180
Qy 181 RSTKHHHTTSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYM 240
Db 181 RSTKHHHTTSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYM 240
Qy 241 LPPSPETGTRTTCILRGYPSSISVQWLPNNEEDHTGHTTTRPKDHGTDPSPFLY 300
Db 241 LPPSPETGTRTTCILRGYPSSISVQWLPNNEEDHTGHTTTRPKDHGTDPSPFLY 300
Qy 301 SRMLVNSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 SRMLVNSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 4

US-09-401-636-9
; Sequence 9, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-9

Query Match 99.0%; Score 1840; DB 10; Length 341;
Best Local Similarity 99.4%; Pred. No. 8.8e-121;
Matches 339; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVW 60
Qy 61 LVDQEAENLFPYTRPRKREGGQTFSLQSEVNITQGMSSNTYTCHVKNGSIFEDSSR 120

Db 61 LVDGQAEANLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
Qy 121 RCDDEPRGVITYLIPSPDLTYNGTPKLCLVLDLSEENITVWVRERKKSIGASQ 180
Db 121 RCDDEPRGVITYLIPSPDLTYNGTPKLCLVLDLSEENITVWVRERKKSIGASQ 180
Qy 181 RSTKHHAATTSITSLPVDADKWIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVYM 240
Db 181 RSTKHNAATTSITSLPVDADKWIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVYM 240
Qy 241 LPPSPETGTRTTCIRGYPSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFFL 300
Db 241 LPPSPETGTRTTCIRGYPSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFFL 300
Qy 301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 5

US-10-176-664-6

; Sequence 6, Application US/10176664

; Publication No. US20030031663A1

; GENERAL INFORMATION:

; APPLICANT: Hellman, Lars T.

; TITLE OF INVENTION: ENHANCED VACCINES

; FILE REFERENCE: 10223/006001

; CURRENT APPLICATION NUMBER: US/10/176,664

; CURRENT FILING DATE: 2002-06-19

; PRIOR APPLICATION NUMBER: US/09/401,636

; PRIOR FILING DATE: 1999-09-22

; PRIOR APPLICATION NUMBER: US 60/106,652

; PRIOR FILING DATE: 1998-11-02

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 341

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetically generated proteins

US-10-176-664-6

Query Match 92.6%; Score 1720; DB 9; Length 341;
Best Local Similarity 93.3%; Pred. No. 2e-112;
Matches 319; Conservative 9; Mismatches 12; Indels 2; Gaps 2;

Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVW 60
Qy 61 LVDGQAEANLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
Db 61 LVDGQAEANLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
Qy 121 RCDDEPRGVITYLIPSPDLTYNGTPKLCLVLDLSEENITVWVRERKKSIGASQ 180
Db 121 RCPDHEPRGVITYLIPSPDLTYNGTAPKLCVLVDLSEKNVNTWVWQEKTSV-NASQ 179
Qy 181 RSTKH-HATTTSITSLPVDADKWIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVY 239
Db 180 WYTKHNNATTSITSLPVAKDWIEGEGYQCIVDHPDFPKPIVRSITKLPKRLAPEVY 239
Qy 240 MLPPSPETGTRTTCIRGYPSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFFL 299
Db 240 MLPPSPETGTRTTCIRGYPSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFFL 299
Qy 300 YSRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 300 YSRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 6

US-09-401-636-6

; Sequence 6, Application US/09401636

; Patent No. US20010038843A1

; GENERAL INFORMATION:

; APPLICANT: Hellman, Lars T.

; TITLE OF INVENTION: ENHANCED VACCINES

; FILE REFERENCE: 10223/006001

; CURRENT APPLICATION NUMBER: US/09/401,636

; CURRENT FILING DATE: 1999-09-22

; PRIOR APPLICATION NUMBER: US 60/106,652

; PRIOR FILING DATE: 1998-11-02

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 341

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetically generated proteins

US-09-401-636-6

Query Match 92.6%; Score 1720; DB 10; Length 341;
Best Local Similarity 93.3%; Pred. No. 2e-112;
Matches 319; Conservative 9; Mismatches 12; Indels 2; Gaps 2;

Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVW 60
Qy 61 LVDGQAEANLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
Db 61 LVDGQAEANLFPYTRPRKREGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
Qy 121 RCDDEPRGVITYLIPSPDLTYNGTPKLCLVLDLSEENITVWVRERKKSIGASQ 180
Db 121 RCPDHEPRGVITYLIPSPDLTYNGAPKLCVLVDLSEKNVNTWVWQEKTSV-NASQ 179
Qy 181 RSTKH-HATTTSITSLPVDADKWIEGEGYQCRVDHPFPKPIVRSITKLPKRLAPEVY 239
Db 180 WYTKHNNATTSITSLPVAKDWIEGEGYQCIVDHPDFPKPIVRSITKLPKRLAPEVY 239
Qy 240 MLPPSPETGTRTTCIRGYPSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFFL 299
Db 240 MLPPSPETGTRTTCIRGYPSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFFL 299
Qy 300 YSRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 300 YSRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 7

US-10-176-664-5

; Sequence 5, Application US/10176664

; Publication No. US20030031663A1

; GENERAL INFORMATION:

; APPLICANT: Hellman, Lars T.

; TITLE OF INVENTION: ENHANCED VACCINES

; FILE REFERENCE: 10223/006001

; CURRENT APPLICATION NUMBER: US/10/176,664

; CURRENT FILING DATE: 2002-06-19

; PRIOR APPLICATION NUMBER: US/09/401,636

; PRIOR FILING DATE: 1999-09-22

; PRIOR APPLICATION NUMBER: US 60/106,652

; PRIOR FILING DATE: 1998-11-02

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 342

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetically generated proteins


```
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-11

Query Match      85.8%; Score 1595; DB 10; Length 341;
Best Local Similarity 85.0%; Pred. No. 1e-103;
Matches 290; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60

Qy 61 LVDQGEAENLPYTRPKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
Db 61 LVDQGEAENLPYTRPKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120

Qy 121 RCDDEPRGVITYLPPSPDLIYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQ 180
Db 121 RCDDEPRGVITYLPPSPDLIYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQ 180

Qy 121 KCSESDPRGVSYLSPPSPDLIYHKA PKITCLVVDLATMEGMNLTWYRESKEPVNPGPL 180
Db 121 KCSESDPRGVSYLSPPSPDLIYHKA PKITCLVVDLATMEGMNLTWYRESKEPVNPGPL 180

Qy 181 RSTKHHTATTSITSLPVDADKMWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYM 240
Db 181 RSTKHHTATTSITSLPVDADKMWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYM 240

Qy 181 NKCHFNGTITVSTLPNTNDWIEGEYTCVTHPLPKDILVRSIAKLPGRKLAPEVYM 240
Db 181 NKCHFNGTITVSTLPNTNDWIEGEYTCVTHPLPKDILVRSIAKLPGRKLAPEVYM 240

Qy 241 LPPSPETGTTTTCVLCIRGFYPSISVQWLPNNEEDHTGHHTTTRPKQKHGTDPSPFFLY 300
Db 241 LPPSPETGTTTTCVLCIRGFYPSISVQWLPNNEEDHTGHHTTTRPKQKHGTDPSPFFLY 300

Qy 301 SRMLVYKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 301 SRMLVYKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

RESULT 11
US-10-176-664-10
; Sequence 10, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-10

Query Match      83.7%; Score 1555; DB 9; Length 345;
Best Local Similarity 84.1%; Pred. No. 6.5e-101;
Matches 290; Conservative 13; Mismatches 38; Indels 4; Gaps 2;

Qy 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60
Db 1 EFHHHHHTLSLPESGPVTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60

Qy 61 LVDQGEAENLPYTRPKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120
Db 61 LVDQGEAENLPYTRPKREGGQTFSLQSEVNITQGMSSNTYTCHVKHNGSIFEDSSR 120

Qy 121 RCDDEPRGVITYLPPSPDLIYENGTPKLTCLVLDLSEENITVWVRERKKSII---G 176
Db 121 RCDDEPRGVITYLPPSPDLIYENGTPKLTCLVLDLSEENITVWVRERKKSII---G 176

Qy 121 RCTAASEPRGVSAVLSPPPTLDLVHKS PKLTCLVVDLASENVNLLWSRENKGGVILPP 180
Db 121 RCTAASEPRGVSAVLSPPPTLDLVHKS PKLTCLVVDLASENVNLLWSRENKGGVILPP 180

Qy 177 SASORSTKHHTATTSITSLPVDADKMWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPE 236
Db 177 SASORSTKHHTATTSITSLPVDADKMWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPE 236

Qy 181 PGPPVIRKPFQNGTFSATSTLFPVNSDMEIETTYCNCVTHPDLPKPILASISKLPGKRLAP 240
Db 181 PGPPVIRKPFQNGTFSATSTLFPVNSDMEIETTYCNCVTHPDLPKPILASISKLPGKRLAP 240

Qy 237 EVMYLPSPSPETGTTTTCVLCIRGFYPSISVQWLPNNEEDHTGHHTTTRPKQKHGTDPSP 296
Db 237 EVMYLPSPSPETGTTTTCVLCIRGFYPSISVQWLPNNEEDHTGHHTTTRPKQKHGTDPSP 296

Qy 241 EVMYLPSPSPETGTTTTCVLCIRGFYPSISVQWLPNNEEDHTGHHTTTRPKQKHGTDPSP 300
Db 241 EVMYLPSPSPETGTTTTCVLCIRGFYPSISVQWLPNNEEDHTGHHTTTRPKQKHGTDPSP 300

Qy 297 FFYLSRMLVYKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
Db 297 FFYLSRMLVYKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

US-10-176-664-8

RESULT 13
US-10-176-664-8
```

```
; Sequence 8, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-8
```

```
Query Match      83.6%; Score 1553.5; DB 9; Length 342;
Best Local Similarity 82.7%; Pred. No. 8.1e-101;
Matches 283; Conservative 21; Mismatches 37; Indels 1; Gaps 1;

Qy 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60
Db 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60

Qy 61 LVDGQAEANLFPYTRPRKREGGQTFSLQSEVNIQTQGMSSNTYTCVKHNGSIFEDSSR 120
Db 61 LVDGQAEANLFPYTRPRKREGGQTFSLQSEVNIQTQGMSSNTYTCVKHNGSIFEDSSR 120

Qy 121 RCDSDPRGVITYLIPSPDLVYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGAS 179
Db 121 RCDSDPRGVITYLIPSPDLVYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGAS 179

Qy 121 KCADSNPRGVSAYLSRSPFDLFIKSPITITCLVVDLAPSKGTVNLWTSRASGKPVNHST 180
Db 121 KCADSNPRGVSAYLSRSPFDLFIKSPITITCLVVDLAPSKGTVNLWTSRASGKPVNHST 180

Qy 180 QRSTKHHHTSITSLPVDAKWIEGEGYOCRVDPHFPKPIVRSITKLPKRLAPEVY 239
Db 180 QRSTKHHHTSITSLPVDAKWIEGEGYOCRVDPHFPKPIVRSITKLPKRLAPEVY 239

Qy 181 RKEEKQRNGTLTVTSLPVGTRDMEGETYQCRVTHPLPALMRSTTKLPKRLAPEVY 240
Db 181 RKEEKQRNGTLTVTSLPVGTRDMEGETYQCRVTHPLPALMRSTTKLPKRLAPEVY 240

Qy 240 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 299
Db 240 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 299

Qy 241 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 300
Db 241 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 300

Qy 300 YSRMLVNKSIWEKGNLVCRVVHEALPGSRITLKSLSHYSAGN 341
Db 301 YSRMLVNKSIWEKGNLVCRVVHEALPGSRITLKSLSHYSAGN 342
```

```
RESULT 14
US-09-401-636-8
; Sequence 8, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-8
```

```
Query Match      83.6%; Score 1553.5; DB 10; Length 342;
Best Local Similarity 82.7%; Pred. No. 8.1e-101;
Matches 283; Conservative 21; Mismatches 37; Indels 1; Gaps 1;

Qy 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60
Db 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60

Qy 61 LVDGQAEANLFPYTRPRKREGGQTFSLQSEVNIQTQGMSSNTYTCVKHNGSIFEDSSR 120
Db 61 LVDGQAEANLFPYTRPRKREGGQTFSLQSEVNIQTQGMSSNTYTCVKHNGSIFEDSSR 120

Qy 121 RCDSDPRGVITYLIPSPDLVYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGAS 179
Db 121 RCDSDPRGVITYLIPSPDLVYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGAS 179

Qy 121 KCADSNPRGVSAYLSRSPFDLFIKSPITITCLVVDLAPSKGTVNLWTSRASGKPVNHST 180
Db 121 KCADSNPRGVSAYLSRSPFDLFIKSPITITCLVVDLAPSKGTVNLWTSRASGKPVNHST 180

Qy 180 QRSTKHHHTSITSLPVDAKWIEGEGYOCRVDPHFPKPIVRSITKLPKRLAPEVY 239
Db 180 QRSTKHHHTSITSLPVDAKWIEGEGYOCRVDPHFPKPIVRSITKLPKRLAPEVY 239

Qy 181 RKEEKQRNGTLTVTSLPVGTRDMEGETYQCRVTHPLPALMRSTTKLPKRLAPEVY 240
Db 181 RKEEKQRNGTLTVTSLPVGTRDMEGETYQCRVTHPLPALMRSTTKLPKRLAPEVY 240

Qy 240 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 299
Db 240 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 299

Qy 241 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 300
Db 241 MLPPSPETGTRTCTLRGFPSEISVQWLNNEDHTGHHTTTRPKQDHDGDPSPFL 300

Qy 300 YSRMLVNKSIWEKGNLVCRVVHEALPGSRITLKSLSHYSAGN 341
Db 301 YSRMLVNKSIWEKGNLVCRVVHEALPGSRITLKSLSHYSAGN 342
```

```
RESULT 15
US-10-176-664-3
; Sequence 3, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-3

Query Match      82.3%; Score 1529; DB 9; Length 341;
Best Local Similarity 80.6%; Pred. No. 4.1e-99;
Matches 275; Conservative 28; Mismatches 38; Indels 0; Gaps 0;

Qy 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60
Db 1 EFHHHHHTLSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHTW 60

Qy 61 LVDGQAEANLFPYTRPRKREGGQTFSLQSEVNIQTQGMSSNTYTCVKHNGSIFEDSSR 120
Db 61 LVDGQAEANLFPYTRPRKREGGQTFSLQSEVNIQTQGMSSNTYTCVKHNGSIFEDSSR 120

Qy 121 RCDSDPRGVITYLIPSPDLVYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGAS 180
Db 121 RCDSDPRGVITYLIPSPDLVYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGAS 180

Qy 181 RSTKHHHTSITSLPVDAKWIEGEGYOCRVDPHFPKPIVRSITKLPKRLAPEVY 240
Db 181 RSTKHHHTSITSLPVDAKWIEGEGYOCRVDPHFPKPIVRSITKLPKRLAPEVY 240

Qy 181 VVKEQYNGTFTVTVSHLPVNTDDWIEGDTYTCRLESPPMVPVPIRTSKAPKRLAPEVY 240
Db 181 VVKEQYNGTFTVTVSHLPVNTDDWIEGDTYTCRLESPPMVPVPIRTSKAPKRLAPEVY 240
```


QY 241 LPPSPETGTTTCTCLIRGFYPSSEISVQWLPNNNEEDHTGHHHTTRPQKDHGTDPSFELY 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 LPPSPETGTTTCTCLIRGFYPSSEISVQWLPNNNEEDHTGHHHTTRPQKDHGTDPSFELY 300
QY 301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 SRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAGN 341

Search completed: July 9, 2003, 13:50:01
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 32 Seconds
(without alignments)
2195.690 Million cell updates/sec

Title: US-09-401-636-4
Perfect score: 1858
Sequence: 1 EFHHHHHTLSLPESGPVTI.....HEALPGSRTEKSLHYSAGN 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 205047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organalle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_xvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488.5	26.3	375	Q9BSZ1	Q9bsz1 homo sapien
2	488.5	26.3	588	Q8WUX4	Q8wux4 homo sapien
3	488.5	26.3	597	Q9BQB8	Q9bqb8 homo sapien
4	488.5	26.3	597	Q96BB9	Q96bb9 homo sapien
5	488.5	26.3	597	Q9BU10	Q9bul10 homo sapien
6	488.5	26.3	613	Q96EY0	Q96ey0 homo sapien
7	488.5	26.3	613	Q8WUK1	Q8wuk1 homo sapien
8	488.5	26.3	614	Q96GAC	Q96gac homo sapien
9	488.5	26.3	618	Q96AA6	Q96aa6 homo sapien
10	463	24.9	337	Q95M34	Q95m34 equus cabal
11	455.5	24.5	613	Q8VCX7	Q8vcx7 mus musculus
12	455	24.5	473	Q8VCX4	Q8vcx4 mus musculus
13	441.5	23.8	471	Q8TC63	Q8tc63 homo sapien
14	435	23.4	437	Q8TC77	Q8tc77 homo sapien
15	434.5	23.4	463	Q9R1A4	Q9r1a4 mus musculus
16	429.5	23.1	469	Q9R1C4	Q9r1c4 mus musculus
				Q8R3V9	Q8r3v9 mus musculus

17	424.5	22.8	473	11	Q9D8L4	Q9d8l4 mus musculus
18	406	21.9	473	11	Q91205	Q91205 mus musculus
19	406	21.9	474	11	Q8R3H6	Q8r3h6 mus musculus
20	400.5	21.6	468	11	Q99L31	Q99l31 mus musculus
21	400.5	21.6	473	11	Q99L25	Q99l25 mus musculus
22	390	21.0	701	4	Q96PQ8	Q96pq8 homo sapien
23	355	19.1	416	4	Q9NPP6	Q9npp6 homo sapien
24	347.5	18.7	384	4	Q9UP60	Q9up60 homo sapien
25	346.5	18.6	494	4	Q96K68	Q96k68 homo sapien
26	346.5	18.6	495	4	Q96KX8	Q96kx8 homo sapien
27	345.5	18.6	495	4	Q96DK0	Q96dk0 homo sapien
28	344.5	18.5	497	4	Q8WY24	Q8wy24 homo sapien
29	336.5	18.1	500	4	Q9BRV0	Q9brv0 homo sapien
30	333.5	17.9	426	11	Q9DCD9	Q9dcd9 mus musculus
31	330.5	17.8	486	11	Q91207	Q91207 mus musculus
32	330.5	17.8	487	11	Q99KA4	Q99ka4 mus musculus
33	326	17.5	479	11	Q99M22	Q99m22 mus musculus
34	326	17.5	481	11	Q8VCV5	Q8vcv5 mus musculus
35	326	17.5	484	11	Q99LA6	Q99la6 mus musculus
36	322.5	17.4	684	13	Q90544	Q90544 ginglymasto
37	314	16.9	488	11	Q91WR1	Q91wr1 mus musculus
38	314	16.9	489	11	Q8VCX4	Q8vcx4 mus musculus
39	313	16.8	481	11	Q91WT3	Q91wt3 mus musculus
40	313	16.8	481	11	Q91WT1	Q91wt1 mus musculus
41	313	16.8	482	11	Q91X92	Q91x92 mus musculus
42	313	16.8	484	11	Q8VEA0	Q8vea0 mus musculus
43	312	16.8	480	11	Q91XE1	Q91xe1 mus musculus
44	311	16.7	479	11	Q91WP5	Q91wp5 mus musculus
45	210	11.3	573	4	Q8WU38	Q8wu38 homo sapien

ALIGNMENTS

RESULT 1

Q9BSZ1 ID Q9BSZ1 PRELIMINARY; PRT; 375 AA.
AC Q9BSZ1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 41.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004476; AA04476.1; -
DR HSP; P01857; IFC1.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003600; IG_like.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 375 AA; 41314 MW; B1A0A0998F473619 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 375;

Best Local Similarity 31.2%; Pred. No. 2.2e-37;

Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

Qy 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSAPKAVHTVWLVGOEAE 69

Db 24 PLPIVTELPKVSVP---VPRDGFNGNPKRS-KLICQATGFSRQIQVSLREGKQVGS 79

Qy 70 LFPVTT-----RPKREGQTFSLSEVNITQGMWSNNTYTVCHKVHNGSIF-EDSSRRCS 123

Db 80 --GVTTDQVQAEKSGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFTQONASSMCV 137
QY 124 DDEPRGVITYLPPSPDLIYENGTKLCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 138 PQDQTAIRVFAIPSPFASIFLTQKLTCLVLDLTYDSVTSWTRQNGEAVKHTNISE 197
QY 184 KHHHATTSITSLPVDADKWIEGEGYQCRVDHPFPKPIVRSITKLPQKRL-APEVYMLP 242
Db 198 SHPNATFSAVGEASICEDDWSNGERFTCTVHTDLPSPKQIISRPGKVALHRPDVYLLP 257
QY 243 PSPEETG--TTRVTCLIRGYPSEISVQWLPNNEEDHTGHTTTRPKDGHGTDSPFLY 300
Db 258 PAREQLNRESATITCLVTGFSADVFVQWMQRGQPLSPKQIISRPGKVALHRPDVYLLP 317
QY 301 SRMLVNKSIWEKGNLVTCTRVVHEALPG---SRTLEKS 334
Db 318 SILTVSEEWNTGTCTVVAHEALPNRVTERTVDKS 354

RESULT 2

Q8WUX4 PRELIMINARY; PRT; 588 AA.
AC Q8WUX4;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical 64.4 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019235; AAH19235.1; -
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG 5.
DR SMART; SM00409; IG 2.
DR SMART; SM00407; IGcl; 4.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 588 AA; 64438 MW; FC60DBAD82B39FD7 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 588;
Best Local Similarity 31.2%; Pred. No. 4.1e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;
QY 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHVTVLVDQGEAEN 69
Db 246 PLPVIAELPPKVSF---VPPRDGFGNPRKS-KLICQATGSPRQIQVSWLRGKQVGS 301
QY 70 LFPYTT-----RPKEGGQTFSLQSEVNITQGWSSNTYTCVKHNGSIF-EDSSRCS 123
Db 302 --GVTTDQVQAEKSGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFTQONASSMCV 359
QY 124 DDEPRGVITYLPPSPDLIYENGTKLCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 360 PQDQTAIRVFAIPSPFASIFLTQKLTCLVLDLTYDSVTSWTRQNGEAVKHTNISE 419
QY 184 KHHHATTSITSLPVDADKWIEGEGYQCRVDHPFPKPIVRSITKLPQKRL-APEVYMLP 242
Db 420 SHPNATFSAVGEASICEDDWSNGERFTCTVHTDLPSPKQIISRPGKVALHRPDVYLLP 479
QY 243 PSPEETG--TTRVTCLIRGYPSEISVQWLPNNEEDHTGHTTTRPKDGHGTDSPFLY 300
Db 480 PAREQLNRESATITCLVTGFSADVFVQWMQRGQPLSPKQIISRPGKVALHRPDVYLLP 539

QY 301 SRMLVNKSIWEKGNLVTCTRVVHEALPG---SRTLEKS 334
Db 540 SILTVSEEWNTGTCTVVAHEALPNRVTERTVDKS 576
RESULT 3
Q9QB08 PRELIMINARY; PRT; 597 AA.
AC Q9QB08;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Unknown (protein for MGC:1905) (protein for MGC:1228).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006180; AAH06180.1; -
DR EMBL; BC001872; AAH01872.1; -
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG cl.
DR InterPro; IPR003600; IG_Like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG 5.
DR SMART; SM00409; IG 2.
DR SMART; SM00407; IGcl; 4.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
SQ SEQUENCE 597 AA; 65300 MW; 2DAFAD50A6375851 CRC64;
Query Match 26.3%; Score 488.5; DB 4; Length 597;
Best Local Similarity 31.2%; Pred. No. 4.2e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;
QY 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHVTVLVDQGEAEN 69
Db 246 PLPVIAELPPKVSF---VPPRDGFGNPRKS-KLICQATGSPRQIQVSWLRGKQVGS 301
QY 70 LFPYTT-----RPKEGGQTFSLQSEVNITQGWSSNTYTCVKHNGSIF-EDSSRCS 123
Db 302 --GVTTDQVQAEKSGPTTYKVTSTLTIKESDWLSQSMFTCRVDHRGLTFTQONASSMCV 359
QY 124 DDEPRGVITYLPPSPDLIYENGTKLCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 360 PQDQTAIRVFAIPSPFASIFLTQKLTCLVLDLTYDSVTSWTRQNGEAVKHTNISE 419
QY 184 KHHHATTSITSLPVDADKWIEGEGYQCRVDHPFPKPIVRSITKLPQKRL-APEVYMLP 242
Db 420 SHPNATFSAVGEASICEDDWSNGERFTCTVHTDLPSPKQIISRPGKVALHRPDVYLLP 479
QY 243 PSPEETG--TTRVTCLIRGYPSEISVQWLPNNEEDHTGHTTTRPKDGHGTDSPFLY 300
Db 480 PAREQLNRESATITCLVTGFSADVFVQWMQRGQPLSPKQIISRPGKVALHRPDVYLLP 539
QY 301 SRMLVNKSIWEKGNLVTCTRVVHEALPG---SRTLEKS 334
Db 540 SILTVSEEWNTGTCTVVAHEALPNRVTERTVDKS 576

RESULT 4

```
Q96BB9
ID Q96BB9 PRELIMINARY; PRT; 597 AA.
AC Q96BB9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 65.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-CELL;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015760; AAH15760.1; -.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGL1; 4.
DR SMART; SM00410; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65039 MW; 4FCA3AD8ECE263D9 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 597;
Best Local Similarity 31.2%; Pred. No. 4.2e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

Qy 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLLCLVSGFSPAKVHVTVLVDGQAEEN 69
Db 246 PLPVAELPPKVSF---VPRDGFNGPRKS-KLICQATGFSRQIQVSLREGKQVGS 301
Qy 70 LFPYTT-----RPKREGQTFSLOSEVNITQGMSSNTYCHVKHNGSIF-EDSSRRCS 123
Db 302 --GVTTDQVQAEKESGPTYKVTSTLTIKESDWLSQSMFCTCRVDHRLGLTFQGNASSMVCV 359
Qy 124 DDEPRGVITYLPPSPDLVYENGTPKLTCLVLVDLESENIITVWVRKKSIGASQSRST 183
Db 360 PQDQTAIRVFAIPSPFASIFLTSTKLTCLVTLDTLYDSVTISWTRQNGEAVKTHTNISE 419
Qy 184 KHHHATTISILPVDADKWIIEGEGYQCRVDHPHPKPIVRSITKLPCKRL-APEVYMLP 242
Db 420 SHPNATFSAVGEASICEDDWNSGERFTCTVTHDLPSPKQITSRPKGVALLRPDVLPLP 479
Qy 243 PSPEETG--TTRTVTLIRGYPSEISVQWLPNNEEDHTGHTTTRPKQDHGTDPSPFLY 300
Db 480 PAREQLNLRSATITCLVTGFSADVFQWMQRGQPLSPKQVTSAPMPEQAPGRYFAH 539
Qy 301 SRMLVNSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
Db 540 SILTVSEEWNTGTYTCVVAHEALPNRVTERTVDKS 576

RESULT 5
Q9BU10 PRELIMINARY; PRT; 597 AA.
ID Q9BU10
AC Q9BU10
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 65.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002963; AAH02963.1; -.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG_c1.
```

```
DR InterPro; IPR003600; IG_like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGL1; 4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65274 MW; 2DAFA8FB7E055851 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 597;
Best Local Similarity 31.2%; Pred. No. 4.2e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

Qy 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLLCLVSGFSPAKVHVTVLVDGQAEEN 69
Db 246 PLPVAELPPKVSF---VPRDGFNGPRKS-KLICQATGFSRQIQVSLREGKQVGS 301
Qy 70 LFPYTT-----RPKREGQTFSLOSEVNITQGMSSNTYCHVKHNGSIF-EDSSRRCS 123
Db 302 --GVTTDQVQAEKESGPTYKVTSTLTIKESDWLSQSMFCTCRVDHRLGLTFQGNASSMVCV 359
Qy 124 DDEPRGVITYLPPSPDLVYENGTPKLTCLVLVDLESENIITVWVRKKSIGASQSRST 183
Db 360 PQDQTAIRVFAIPSPFASIFLTSTKLTCLVTLDTLYDSVTISWTRQNGEAVKTHTNISE 419
Qy 184 KHHHATTISILPVDADKWIIEGEGYQCRVDHPHPKPIVRSITKLPCKRL-APEVYMLP 242
Db 420 SHPNATFSAVGEASICEDDWNSGERFTCTVTHDLPSPKQITSRPKGVALLRPDVLPLP 479
Qy 243 PSPEETG--TTRTVTLIRGYPSEISVQWLPNNEEDHTGHTTTRPKQDHGTDPSPFLY 300
Db 480 PAREQLNLRSATITCLVTGFSADVFQWMQRGQPLSPKQVTSAPMPEQAPGRYFAH 539
Qy 301 SRMLVNSIWEKGNLVTCRVVHEALPG---SRTLEKS 334
Db 540 SILTVSEEWNTGTYTCVVAHEALPNRVTERTVDKS 576

RESULT 6
Q96EY0 PRELIMINARY; PRT; 613 AA.
ID Q96EY0
AC Q96EY0
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:20337).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-CELL;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011857; AAH11857.1; -.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; ig; 5.
DR SMART; SM00408; IGV; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Immunoglobulin domain.
SQ SEQUENCE 613 AA; 67273 MW; 31214203FB8421E7 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 613;
Best Local Similarity 31.2%; Pred. No. 4.3e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

Qy 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLLCLVSGFSPAKVHVTVLVDGQAEEN 69
```

Db 241 PLPVIAELPPKVSF---VPRDGFNPKS-KLICQATGSPRQIQVSWLRGKQVGS 296
Qy 70 LPPYTT-----RPKEGGQTSLOSEVNITQGMSSNTYCHVKHNSIF-EDSSRCS 123
Db 297 --GVTTDQVQAEKESGPTTKVSTLTIKESDWLSQSMFTCRVDHRLTFOQNASSMCV 354
Qy 124 DDEPRGVITYLPPSPDLNGLYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 355 PQDQTAIRVFAIPSPFASIFLTKSTKLTCLVTLDTYDSVTISWTRNGEAVKTHNISE 414
Qy 184 KHHHATTSTISILPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APEVYMLP 242
Db 415 SHPNATFSAVGEASICEDDWNSSGERFTCTVTHDPLSPKLTISRPKGVALLHRPDVYLLP 474
Qy 243 PSPEETG--TTRVTCLIRGFYSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFPLY 300
Db 475 PAREQLNRESATITCLVTGFSPADVFQWQMGQPLSPKQVTSAPMPBPQAPGRYFAH 534
Qy 301 SRMLVNSIWEKGNLVTCTVVAHEALPG---SRTLEKS 334
Db 535 SILTVSEEWNTGETYTCVVAHEALPNRVTERTVDKS 571

RESULT 7

Q8WUK1
ID Q8WUK1 PRELIMINARY; PRT; 613 AA.
AC Q8WUK1
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 67.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020240; AAH02040.1; -
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG.c1.
DR InterPro; IPR003006; IG.MHC.
DR InterPro; IPR003596; IG.V.
DR Pfam; PF00047; IG.5.
DR SMART; SM00409; IG.2.
DR SMART; SM00407; IGc1; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG.MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671B315 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 613;
Best Local Similarity 31.2%; Pred. No. 4.3e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;
Qy 17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLCLVSGSPAKVHVTLVDGQBAEN 69
Db 241 PLPVIAELPPKVSF---VPRDGFNPKS-KLICQATGSPRQIQVSWLRGKQVGS 296
Qy 70 LPPYTT-----RPKEGGQTSLOSEVNITQGMSSNTYCHVKHNSIF-EDSSRCS 123
Db 297 --GVTTDQVQAEKESGPTTKVSTLTIKESDWLSQSMFTCRVDHRLTFOQNASSMCV 354
Qy 124 DDEPRGVITYLPPSPDLNGLYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 355 PQDQTAIRVFAIPSPFASIFLTKSTKLTCLVTLDTYDSVTISWTRNGEAVKTHNISE 414
Qy 184 KHHHATTSTISILPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APEVYMLP 242
Db 415 SHPNATFSAVGEASICEDDWNSSGERFTCTVTHDPLSPKLTISRPKGVALLHRPDVYLLP 474

Query Match 26.3%; Score 488.5; DB 4; Length 613;
Best Local Similarity 31.2%; Pred. No. 4.3e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;
Qy 17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLCLVSGSPAKVHVTLVDGQBAEN 69
Db 241 PLPVIAELPPKVSF---VPRDGFNPKS-KLICQATGSPRQIQVSWLRGKQVGS 296
Qy 70 LPPYTT-----RPKEGGQTSLOSEVNITQGMSSNTYCHVKHNSIF-EDSSRCS 123
Db 297 --GVTTDQVQAEKESGPTTKVSTLTIKESDWLSQSMFTCRVDHRLTFOQNASSMCV 354
Qy 124 DDEPRGVITYLPPSPDLNGLYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 355 PQDQTAIRVFAIPSPFASIFLTKSTKLTCLVTLDTYDSVTISWTRNGEAVKTHNISE 414
Qy 184 KHHHATTSTISILPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APEVYMLP 242
Db 415 SHPNATFSAVGEASICEDDWNSSGERFTCTVTHDPLSPKLTISRPKGVALLHRPDVYLLP 474

Qy 243 PSPEETG--TTRVTCLIRGFYSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFPLY 300
Db 475 PAREQLNRESATITCLVTGFSPADVFQWQMGQPLSPKQVTSAPMPBPQAPGRYFAH 534
Qy 301 SRMLVNSIWEKGNLVTCTVVAHEALPG---SRTLEKS 334
Db 535 SILTVSEEWNTGETYTCVVAHEALPNRVTERTVDKS 571
RESULT 8
Q96GA6
ID Q96GA6 PRELIMINARY; PRT; 614 AA.
AC Q96GA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:15420).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=B-CELL;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009851; AAH09851.1; -
DR InterPro; IPR000005; HTHARAC.
DR InterPro; IPR003598; IG.C2.
DR InterPro; IPR003006; IG.MHC.
DR Pfam; PF00047; IG.5.
DR SMART; SM00408; IGc2; 2.
DR PROSITE; PS00041; HTH ARAC FAMILY 1; UNKNOWN_1.
DR PROSITE; PS00290; IG.MHC; UNKNOWN_3.
KW Immunoglobulin domain.
SQ SEQUENCE 614 AA; 67921 MW; 55EF536E77AA9BBB CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 614;
Best Local Similarity 31.2%; Pred. No. 4.3e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;
Qy 17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLCLVSGSPAKVHVTLVDGQBAEN 69
Db 242 PLPVIAELPPKVSF---VPRDGFNPKS-KLICQATGSPRQIQVSWLRGKQVGS 297
Qy 70 LPPYTT-----RPKEGGQTSLOSEVNITQGMSSNTYCHVKHNSIF-EDSSRCS 123
Db 298 --GVTTDQVQAEKESGPTTKVSTLTIKESDWLSQSMFTCRVDHRLTFOQNASSMCV 355
Qy 124 DDEPRGVITYLPPSPDLNGLYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 356 PQDQTAIRVFAIPSPFASIFLTKSTKLTCLVTLDTYDSVTISWTRNGEAVKTHNISE 415
Qy 184 KHHHATTSTISILPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APEVYMLP 242
Db 416 SHPNATFSAVGEASICEDDWNSSGERFTCTVTHDPLSPKLTISRPKGVALLHRPDVYLLP 475
Qy 243 PSPEETG--TTRVTCLIRGFYSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFPLY 300
Db 476 PAREQLNRESATITCLVTGFSPADVFQWQMGQPLSPKQVTSAPMPBPQAPGRYFAH 535
Qy 301 SRMLVNSIWEKGNLVTCTVVAHEALPG---SRTLEKS 334
Db 536 SILTVSEEWNTGETYTCVVAHEALPNRVTERTVDKS 572

Query Match 26.3%; Score 488.5; DB 4; Length 614;
Best Local Similarity 31.2%; Pred. No. 4.3e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;
Qy 17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLCLVSGSPAKVHVTLVDGQBAEN 69
Db 242 PLPVIAELPPKVSF---VPRDGFNPKS-KLICQATGSPRQIQVSWLRGKQVGS 297
Qy 70 LPPYTT-----RPKEGGQTSLOSEVNITQGMSSNTYCHVKHNSIF-EDSSRCS 123
Db 298 --GVTTDQVQAEKESGPTTKVSTLTIKESDWLSQSMFTCRVDHRLTFOQNASSMCV 355
Qy 124 DDEPRGVITYLPPSPDLNGLYENGTPKLTCLVLDLSEENITVWVRERKKSIGASQSRST 183
Db 356 PQDQTAIRVFAIPSPFASIFLTKSTKLTCLVTLDTYDSVTISWTRNGEAVKTHNISE 415
Qy 184 KHHHATTSTISILPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APEVYMLP 242
Db 416 SHPNATFSAVGEASICEDDWNSSGERFTCTVTHDPLSPKLTISRPKGVALLHRPDVYLLP 475
Qy 243 PSPEETG--TTRVTCLIRGFYSEISVQWLPNNEEDHTGHTTTRPKQKHGTDPSFPLY 300
Db 476 PAREQLNRESATITCLVTGFSPADVFQWQMGQPLSPKQVTSAPMPBPQAPGRYFAH 535
Qy 301 SRMLVNSIWEKGNLVTCTVVAHEALPG---SRTLEKS 334
Db 536 SILTVSEEWNTGETYTCVVAHEALPNRVTERTVDKS 572

RESULT 9
Q96AA6
ID Q96AA6 PRELIMINARY; PRT; 618 AA.
AC Q96AA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypothetical 67.8 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017356; AAH17356.1; -.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG_5.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 618 AA; 67758 MW; 96BBD4C7C696E0A6 CRC64;

Query Match 26.3%; Score 488.5; DB 4; Length 618;
Best Local Similarity 31.2%; Pred. No. 4.4e-37;
Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

QY 17 PVTII---PPTVKLFHSSCDPR----GDAHSTIQLCLVSGFSPAKVHVTVLVDQAEAN 69
DB 246 PLPVAIEALPPKVSF---VPRDGFPGNPKRS-KLICQATGFSRQIQVSWLREGKQVGS 301
QY 70 LPPVTT----RPKREGGQTSLOSEVNITOGQWMSNTYCHVKHNGSIF-EDSSRCS 123
DB 302 -GVTTDQVQAEKESGPTTKVSTLTIKESDMLSQSMFTCRVDHRLTQQNASSMCV 359
QY 124 DDEPRGVTIYLPPSPDLNYPKLTCLVLDLSEENITVTVRERKKSIGSASQST 183
DB 360 PDQDTAIRVFAIPSPSFASIFLTSTKLTCLVTLDTLYDSVITSMTQNGEAVKHTNISE 419
QY 184 KHHATTSTIILPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYMLP 242
DB 420 SHPNATFSVAGEASICEEDMNSGERFTCTVTHTLPSLPKQTSIRPKGVALHPRDVPYLLP 479
QY 243 PSPEETG--TTRVTCLIRGFVPSISVQWLPNNEEDTGHHTTTRPKDGHGTDPSRFLY 300
DB 480 PARQLNRESATITCLVTGSPADVFQWQMGQPLSPKQVTSAPMPFQAPGRYPFAH 539
QY 301 SRMLVNSIKWKNLVTCTRVVHEALPG---SRTLEKS 334
DB 540 SILTVSEEWNTGETYTTCVAHEALPNRVTRTVDKS 576

RESULT 10
Q95M34 PRELIMINARY; PRT; 337 AA.
AC Q95M34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
DE (Fragment).
GN IGHCl.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner B.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.,
RT "Organization of the equine immunoglobulin heavy chain constant region
genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";

RL Immunobiology 199:105-119(1998).
DR EMBL; AJ300675; CAC44624.1; -.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG_2; MHC; UNKNOWN_2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 24.9%; Score 463; DB 6; Length 337;
Best Local Similarity 31.0%; Pred. No. 4.7e-35;
Matches 107; Conservative 71; Mismatches 121; Indels 46; Gaps 12;

QY 23 PTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVTVM----LVDGQAEANLFFPYTRPK 78
DB 6 PKVFALAPGCGTSD--STVALGCLVSGYFPEPVKVSWSGSLTSG---VHTFFSVL--- 57
QY 79 REGQTSLOSEVNITOGQWMSNTYCHVKHNGSIFEDSSR----- 120
DB 58 -QSSGFYSLSMVTVPASTW-TSETYICNVVHAASNFVKDKRIBPDPNHQKVCDSKCP 115
QY 121 RCDDE-PRGVITVLIIPSPDL-LYENGTPKLTCLVLDLSEENITVTVRERKKSIGS 177
DB 116 KCPAPELLGSPSVIFPPNPKDTLMTITREVTICVVDVSGQENPDVKNWMDGVEVRTA 175
QY 178 ASQRSTKHHTTSTIILPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPE 237
DB 176 TTRPKEQFNSTYRVVSVLRIGHQDWLSGKEFKCKVNNQALPQPIERTITTKGRSQEPQ 235
QY 238 VYMLPPSPETGYTR-TVTCIRGFYSEISVQWLPNNEEDTGHHTTTRPKDGHGTDPS 296
DB 236 VYVLAPHDELKSKVSTCLVKDFYFPEINIEQSNQSQPELETKYSTQAQOD--SDGS 293
QY 297 FFYLSRLVNKSTWKNLVTCTRVVHEALPGSRTLEKSLHVSAGN 341
DB 294 YFLYSLKSVDRNRMQQTTFCTTCGVWHEALHN-----HTQKN 330

RESULT 11
Q8VCX7 PRELIMINARY; PRT; 613 AA.
AC Q8VCX7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 67.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SALIVARY GLAND;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018315; AAH18315.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; IG_5.
DR SMART; SM00409; IG_2.
DR SMART; SM00407; IGC1; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67855 MW; 41A9384DD4C22862 CRC64;

Query Match 24.5%; Score 455.5; DB 11; Length 613;
Best Local Similarity 32.1%; Pred. No. 5.3e-34;
Matches 105; Conservative 63; Mismatches 140; Indels 19; Gaps 8;

QY 23 PTVKLFHSSCDPR----GDAHSTIQLCLVSGFSPAKVHVTVLVDGQAEANLFF---PYTT 75

```
Db 249 PNVNVF---VPRDQGFSGAPRKSLICEATNFTPKITVSWLKGDLVSGFTTDPVTI 305
Qy 76 RPKEGGQTFSLQSVNITQGMSSNTYTHVKNHSIF-EDSRRCSDDPRGVITYL 134
Db 306 ENKSGTPQYKIVISTLTISEDLNLTNYTKRVDHRLGTLTKNVSSTCAASPDLITFT 365
Qy 135 IPPSPDLIYENGTPKLTCLVLDLESEENITVWVRERKKSIGSASQSRQTKHHHTSITS 194
Db 366 IPPSPADIFLSKANLTCLVSNLATYETLNISWASQSGEPLTKIKIMESHNGTFFSAKG 425
Qy 195 ILPVDKAWIEGEGYQCRVDHPHFKPIVRSITKLPKGK-RIAPVYMLPPSPBETG--T 250
Db 426 VASVCVEDWNRKKEFVCTVTRDLSPQKFTSK-PNEVHKHPPAVYLLPPAREQLNIRE 484
Qy 251 TRVTCLIRGFVPSRISVQWLPNNEEDHTGHHTTRPKQDHTDPSFFLYSRMLVKNKSIW 310
Db 485 SATVTCVLKGFSPADISVQWLRQGLLPQEKVTSAPMPGAPGFYTHSILTVTEEW 544
Qy 311 EKGNIJVTCTVGVHEALP---GSRTELEKS 334
Db 545 NSGETYTCVGVHEALPHLVTRTVDKS 571

RESULT 12
Q8TC63
ID Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025985; AAH5985.1; -.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 24.5%; Score 455; DB 4; Length 473;
Best Local Similarity 33.4%; Pred. No. 4.2e-34;
Matches 115; Conservative 67; Mismatches 128; Indels 34; Gaps 12;

Qy 18 VTIIPTVK---LPHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTV---LVDGQAEAL 70
Db 142 VSVSPASTKGPSVFPPLAPCSRSTSSSTAALGCLVKDYFPEPTVSWNSGALTSG--VHT 198
Qy 71 PPTTRPKREGGQTFSLQSEVNITQGMSSNTYTHVKNHSIFEDSSRCSDDPE---127
Db 199 PPAVL-----QSSGLYSLSVWTVPPSS-LGTQTYICNVNHNKPSNTKVDKVEKSGPPCP 253
Qy 128 -----RGVITYILIPSPDL-LYENGTPKLTCLVLDLESEE-NITVWVRERKKSIGS 177
Db 254 SCPAPEFLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSDDEPVEVQNVVDVGVVNA 313
Qy 178 ASQSTKHHHTSITSLPVDKADWIEGEGYQCRVDHPHFKPIVRSITKLPKRLAPE 237
Db 314 KTKPREEQFNSTYRVVSVLTVLHQDLNGKEYCKVSNKGLPSSIEKTSKAKGQPREPQ 373
Qy 238 VYMLPPSPBEE-TGTRTTCVIRGFYPSSEISVQWLPNNEEDHTGHHTTRPKQDHTDPS 296
Db 374 VYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN--NYKTTTPVLD--SDGS 429
Qy 297 FFYLSRMLVKNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
Db 430 FFYLSRLTVDKSRWQEGNVFSCSVNHEALHNHYT-QKSLSLSLG 472
```

```
RESULT 13
Q8TC77
ID Q8TC77 PRELIMINARY; PRT; 471 AA.
AC Q8TC77
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 51.8 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024289; AAH24289.1; -.
KW Hypothetical protein.
SQ SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;

Query Match 23.8%; Score 441.5; DB 4; Length 471;
Best Local Similarity 31.7%; Pred. No. 7.6e-33;
Matches 110; Conservative 73; Mismatches 127; Indels 37; Gaps 13;

Qy 18 VTIIPTVK---LPHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTV---LVDGQAEAL 70
Db 137 VTSSASTKGPSVFPPLAPSKSTSGTAALGCLVKDYFPEPTVSWNSGALTSG--VHT 193
Qy 71 PPTTRPKREGGQTFSLQSEVNITQGMSSNTYTHVKNHSIFE-----DSSR 120
Db 194 PPAVL-----QSSGLYSLSVWTVPPSS-LGTQTYICNVNHNKPSNTKVDKVEKSCDKTH 248
Qy 121 RC-----SDDEPRGVITYILIPSPDL-LYENGTPKLTCLVLDLESEE-NITVWVRERKKS 174
Db 249 TCPECPAPELGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHDEPVEVQNVVDGVVEV 308
Qy 175 IGSASORSTKHHHTSITSLPVDKADWIEGEGYQCRVDHPHFKPIVRSITKLPKRL 234
Db 309 HNATKPREEQFNSTYRVVSVLTVLHQDLNGKEYCKVSNKALPALEKTSKAKGQPR 368
Qy 235 APEVYMLPPSPBEE-TGTRTTCVIRGFYPSSEISVQWLPNNEEDHTGHHTTRPKQDHTG 293
Db 369 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN--NYKTTTPVLD--S 424
Qy 294 DPSEFLYSRMLVKNKSIWEKGNLVTCRVVHEALPGSRTLEKSLHYSAG 340
Db 425 DGSFFLYSKLTVDKSRWQEGNVFSCSVNHEALHNHYT-QKSLSLSPG 470

RESULT 14
Q9RI14
ID Q9RI14 PRELIMINARY; PRT; 437 AA.
AC Q9RI14
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gamma1 heavy chain of Mab7 (Fragment).
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G.; Yu X.; Ekramoddoullah A.K.M.; Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal antibody (Mab 7, its light and heavy chains) and construction of a single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -.
DR HSPF; F01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
```


DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; Ig_v. 1.
DR SMART; SM00410; Ig_like. 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EB7D697C CRC64;

Query Match 23.4%; Score 435; DB 11; Length 437;
Best Local Similarity 31.2%; Pred. No. 2.8e-32;
Matches 109; Conservative 66; Mismatches 134; Indels 40; Gaps 13;

Qy 16 GPVTII-----PPTVKLFHSSCDPRGDAHSTIQLCLYSGFSPAKVHTW---LVD 63
Db 104 GPGIUVTSAAKTTSPSV--YPLAGSAAQTNSMWTGLCLVKGYPPEVVTWNSGSLSS 161
Qy 64 QGEAENLFPYTRPKREGGQTFSLQSEVNIQTGQWMSNTYTCHVKHNGSIF----- 115
Db 162 G---VHTFPAVLQ-----SDLYTLSSSVTPSSTW-PSETVTCNVAHPASSTKVDDKI 212
Qy 116 -EDSRRRCSDDEPRGVIITLIPPSPLD-LYENGTPKLTCLVLDLESSE-NITVTWVRERK 172
Db 213 RDCGCKPICITVPEVSSVFIFFPKPKDVLITLTPEKVTCCVVVDISKDDPEVQFSWFVDDV 272
Qy 173 KSIGSASQSRSTKHHTTITSILPVDADKMWIEGEGYQCRVDHPHPKPIVRSITKLPGK 232
Db 273 EVHTAQTPREEQFNSTFRSVSELPIMHQDMLNGKEFKCRVNSAAPPAPIEKTISKTKGR 332
Qy 233 RLAPVYMLPPSPETGTTR-TVTCLIRGFVPSEISVQWLPNNEEDHTGHHTTTRPQKDH 291
Db 333 PKAPQVYTIPTPKQEQMAKDKVSLTCTMTDFPEDITVQWNGQP--AENYKNTQPI 389
Qy 292 GTDPSFFLYSRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAG 340
Db 390 -TDGSYFYIYKLVNQKSNWEAGNTFTCSVLHGLHNHHT-EKSLSHSPG 436

RESULT 15
Q99LC4 PRELIMINARY; PRT; 463 AA.
AC Q99LC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AA003435.1; --
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 2.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BEC30783 CRC64;

Query Match 23.4%; Score 434.5; DB 11; Length 463;
Best Local Similarity 31.6%; Pred. No. 3.4e-32;
Matches 110; Conservative 65; Mismatches 136; Indels 37; Gaps 13;

Qy 9 TSLPESGPTIIPPTVKLFHSSCDPRGDAHSTIQLCLYSGFSPAKVHTW---LVDG 64
Db 136 TVSAAKTTPPSVYP---LAPGSA---AQTNMWTGLCLVKGYPPEVVTWNSGSLSSG 188
Qy 65 QGEAENLFPYTRPKREGGQTFSLQSEVNIQTGQWMSNTYTCHVKHNGSIF----- 115
Db 189 ---VHTFPAVLQ-----SDLYTLSSSVTPSSTW-PSETVTCNVAHPASSTKVDDKI 239
Qy 116 EDSRRRCSDDEPRGVIITLIPPSPLD-LYENGTPKLTCLVLDLESSE-NITVTWVRERK 173
Db 240 DCGCKPICITVPEVSSVFIFFPKPKDVLITLTPEKVTCCVVVDISKDDPEVQFSWFVDDVE 299
Qy 174 SIGSASQSRSTKHHTTITSILPVDADKMWIEGEGYQCRVDHPHPKPIVRSITKLPGKR 233
Db 300 VHTAQTPREEQFNSTFRSVSELPIMHQDMLNGKEFKCRVNSAAPPAPIEKTISKTKGR 359
Qy 234 LAPVYMLPPSPETGTTR-TVTCLIRGFVPSEISVQWLPNNEEDHTGHHTTTRPQKDH 292
Db 360 KAPQVYTIPTPKQEQMAKDKVSLTCTMTDFPEDITVQWNGQP--AENYKNTQPI 415
Qy 293 TDPSPFFLYSRMLVNKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAG 340
Db 416 TDGSYFYIYKLVNQKSNWEAGNTFTCSVLHGLHNHHT-EKSLSHSPG 462

Search completed: July 9, 2003, 13:41:03
Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 11 Seconds
(without alignments)
1285.766 Million cell updates/sec

Title: US-09-401-636-4

Perfect score: 1858

Sequence: 1 EFHHHHHTLSLPESGPVTI.....HEALPGSRITLKSLLHYSAGN 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1051	56.6	429	1 EPC RAT	P01855 rattus norv
2	940.5	50.6	421	1 EPC_MOUSE	P03336 mus musculus
3	745	40.1	428	1 EPC_HUMAN	P01854 homo sapien
4	499	26.9	454	1 MUC_HUMAN	P01871 homo sapien
5	490	26.4	479	1 MUCM RABIT	P04221 oryctolagus
6	488	26.3	458	1 MUC RABIT	P03988 oryctolagus
7	484.5	26.1	391	1 MUCB_HUMAN	P04220 homo sapien
8	468	25.2	326	1 GC2_HUMAN	P01859 homo sapien
9	464.5	25.0	450	1 MUC_CANFA	P01874 canis fami
10	455.5	24.5	455	1 MUC_MOUSE	P01872 mus musculus
11	455.5	24.5	476	1 MUCM_MOUSE	P01873 mus musculus
12	451.5	24.3	327	1 GC4_HUMAN	P01861 homo sapien
13	449.5	24.2	457	1 MUC_SUNMU	P02768 suncus muri
14	440.5	23.7	330	1 GC1_HUMAN	P01857 homo sapien
15	433.5	23.3	454	1 MUC_MESAU	P06337 mesocricetu
16	426	22.9	322	1 GCA RAT	P02760 rattus norv
17	424.5	22.8	325	1 GGC RAT	P02762 rattus norv
18	424.5	22.8	335	1 GCAE_MOUSE	P01864 mus musculus
19	423	22.8	326	1 GC1 RAT	P02759 rattus norv
20	422	22.7	323	1 GC_RABIT	P01870 oryctolagus
21	419	22.6	329	1 GC2_CAVPO	P01862 cavia porce
22	410	22.1	324	1 GC1_MOUSE	P01868 mus musculus
23	410	22.1	393	1 GC1M_MOUSE	P01869 mus musculus
24	408.5	22.0	329	1 GC3_MOUSE	P22436 mus musculus
25	405	21.8	336	1 GCB_MOUSE	P01866 mus musculus
26	405	21.8	405	1 GCBM_MOUSE	P01867 mus musculus
27	403.5	21.7	398	1 GC3M_MOUSE	P03987 mus musculus
28	394	21.2	330	1 GC3A_MOUSE	P01863 mus musculus
29	394	21.2	399	1 GCAM_MOUSE	P01865 mus musculus
30	386	20.8	290	1 GC3_HUMAN	P01860 homo sapien
31	386	20.8	438	1 HVC2_HETFR	P23085 heterodontu
32	381.5	20.5	333	1 GCB RAT	P02761 rattus norv
33	373	20.1	438	1 HVC5_HETFR	P23087 heterodontu

RESULT 1

ID	EPC RAT	STANDARD	PRT	429 AA
AC	P01855;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	ig epsilon chain C region.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthesia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).			
RC	STRAIN=LOU/C/WSL;			
RX	MEDLINE=83064537; PubMed=6292865;			
RA	Hellman L., Petterson U., Engstroem A., Karlsson T., Bennich H.;			
RT	"Structure and evolution of the heavy chain from rat immunoglobulin E.";			
RT	Nucleic Acids Res. 10:6041-6049(1982).			
RN	[2]			
RP	SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).			
RX	MEDLINE=83182019; PubMed=6820340;			
RA	Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;			
RT	"A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction, identification, and DNA sequence.";			
RL	DNA 1:335-343(1982).			
RN	[3]			
RP	SEQUENCE OF 205-306 FROM N.A.			
RX	MEDLINE=82174576; PubMed=6803238;			
RA	Hellman L., Petterson U., Bennich H.;			
RT	"Characterization and molecular cloning of the mRNA for the heavy (epsilon) chain of rat immunoglobulin E.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; J00744; AAA41379.1; ALT_INIT.			
DR	PIR; A02143; EHRT.			
DR	HSSP; P01854; 1IGE.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003597; Ig_C1.			
DR	InterPro; IPR003600; Ig_like.			
DR	Pfam; PF00047; Ig; 4.			
DR	SMART; SM00410; IG like; 3.			
DR	SMART; SM00407; IGE1; 1.			
DR	PROSITE; PS00290; IG_MHC; 3.			
KW	Immunoglobulin domain; Immunoglobulin C region.			
FT	NON TER 1			
FT	CONFLICT 168 168 R -> N (IN REF. 2).			

ALIGNMENTS

34	373	20.1	461	1	HVCM_HETFR	P23088 heterodontu
35	363	19.5	393	1	HVC3_HETFR	P23086 heterodontu
36	355	19.1	340	1	ALC2_HUMAN	P01877 homo sapien
37	348	18.7	370	1	HVC1_HETFR	P23084 heterodontu
38	346.5	18.6	353	1	ALC1_HUMAN	P01876 homo sapien
39	336.5	18.1	353	1	ALC1_GORGO	P20758 gorilla gor
40	334.5	18.0	446	1	MUC_CHICK	P01875 gallus gall
41	323	17.4	344	1	ALC_MOUSE	P01878 mus musculus
42	270.5	14.6	481	1	MUCM ICTPU	P23735 ictalurus p
43	237.5	12.8	299	1	ALC_RABIT	P01879 oryctolagus
44	210	11.3	383	1	DTC_HUMAN	P01880 homo sapien
45	160	8.6	103	1	LAC_CHICK	P20763 gallus gall

RN SEQUENCE FROM N.A.
 RP MEDLINE=83168897; PubMed=6300763;
 RA Seno M., Kurokawa T., Ono Y., Onda H., Saeada R., Igarashi K.,
 RA Kikuchi M., Sugino Y., Nishida Y., Honjo T.;
 RT "Molecular cloning and nucleotide sequencing of human immunoglobulin
 RT epsilon chain cDNA.";
 RL Nucleic Acids Res. 11:719-726(1983).
 [2]
 RN SEQUENCE FROM N.A., AND VARIANT LEU-359.
 RP MEDLINE=83001945; PubMed=6288268;
 RA Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;
 RT "Duplication and deletion in the human immunoglobulin epsilon genes.";
 RL Cell 29:691-699(1982).
 [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=84236029; PubMed=6234164;
 RA Planagan J.G., Rabbitts T.H.;
 RT "The sequence of a human immunoglobulin epsilon heavy chain constant
 RT region gene, and evidence for three non-allelic genes.";
 RL EMBO J. 1:655-660(1982).
 [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=84207910; PubMed=6327276;
 RA Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;
 RT "Long terminal repeat-like elements flank a human immunoglobulin
 RT epsilon pseudogene that lacks introns.";
 RL EMBO J. 1:1539-1544(1982).
 [5]
 RN PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).
 RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;
 RL (In) Bach M.K. (eds.);
 RT Immediate hypersensitivity: modern concepts and developments, pp.1-36,
 RL Marcel Dekker, New York (1978).
 [6]
 RN SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.
 RP MEDLINE=83065234; PubMed=6815656;
 RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,
 RA Bell L.O., Gould H.J.;
 RT "Cloning and sequence determination of the gene for the human
 RT immunoglobulin epsilon chain expressed in a myeloma cell line.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).
 [7]
 RN 3D-STRUCTURE MODELING.
 RP MEDLINE=87089848; PubMed=3796618;
 RA Padlan E.A., Davies D.R.;
 RT "A model of the Fc of immunoglobulin E.";
 RL Mol. Immunol. 23:1063-1075(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L00022; AAB59424.1; ALT_INIT.
 DR PIR; A02142; EHHU.
 DR PIR; A22771; A22771.
 DR PIR; A23195; A23195.
 DR PDB; 1TGE; 15-JUL-92.
 DR Genew; HGNC:5522; IGHE.
 DR MIM; 147180; -.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003597; IG_c1.
 DR Pfam; PF00047; IG; 4.
 DR SMART; SM00407; IGc1; 4.
 DR PROSITE; PS00290; IG_MHC; 3.
 DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW 3D-structure. 1 14
 FT NON_TER 1 14
 FT DISULFID 14 INTERCHAIN (WITH A LIGHT CHAIN).

FT DISULFID 15 105
 FT DISULFID 29 85
 FT DISULFID 121 121
 FT DISULFID 135 193
 FT DISULFID 209 209
 FT DISULFID 239 299
 FT DISULFID 345 405
 FT CARBOHYD 21 21
 FT CARBOHYD 49 49
 FT CARBOHYD 99 99
 FT CARBOHYD 146 146
 FT CARBOHYD 252 252
 FT CARBOHYD 275 275
 FT VARIANT 359 359
 FT /FTid=VAR_003885.
 SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;
 Query Match 40.1%; Score 745; DB 1; Length 428;
 Best Local Similarity 44.4%; Pred. No. 3.9e-49;
 Matches 143; Conservative 62; Mismatches 109; Indels 8; Gaps 5;
 QY 22 PPTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVHTLVLDGQAEANLFFYTRPKREG 81
 DB 111 PPTVKILQSSCDGGHPPPTIQLCLVSGYTPGTINITWLEDGQVMD--VDLSTASTQE 168
 QY 82 QTFSLQSEVNITQGMWSNNTYCHVKHNGSIFEDSSRRCSDDPRGVITYLIPPSPLD 141
 DB 169 GELASTQSELTLSQKHWLSDRTYTCQVYQHTFEDSTKKCADSNPRGVSAYLSRPSFD 228
 QY 142 LYENGTPKLCVLVDL--ESEENITVWVRKKSGISASORSTKHHTATTSITILPVD 200
 DB 229 LFIRKSPITICLVVDLAPSGTGNLTWSRASGKPVNHSRKEQRNGTLTVSTLPVGT 288
 QY 201 KDWTEGSGYOCRVDPHPPKPIVRSITKLPCKRLAPVYMLPPSPETGT--TETVTCLI 258
 DB 289 RDMIEGETYQCRVTHPLPALMKSTTKTSGRAPEVIAP-APEWPGSKDKETLACLI 347
 QY 259 RGFPSPISQWLPNNEEDTHGHTTTRPQKHGTDPSFFLYSLMVKNSIWEKGNLVT 318
 DB 348 QNFNPEDISQWLHNEVQLPDARHSTTPRKTKS--GFFVFSRLVETRAWEQKDEFIC 405
 QY 319 RVVHEALPGSRTEKSLHYSAG 340
 DB 406 RAVHEAASPSQTVQRAVSNVPG 427
 RESULT 4
 MUC HUMAN
 ID MUC HUMAN STANDARD; PRT; 454 AA.
 AC P01871;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ig mu chain C region.
 GN IGHM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OK NCBI_Taxid=9606;
 [1]
 RN SEQUENCE OF 1-434 FROM N.A.
 RP MEDLINE=90332450; PubMed=2115996;
 RA Friedlander R.M., Nussenzweig M.C., Leder P.;
 RT "Complete nucleotide sequence of the membrane form of the human IGM
 RT heavy chain.";
 RL Nucleic Acids Res. 18:4278-4278(1990).
 [2]
 RN SEQUENCE (WALDENSTROM'S MACROGLOBULIN GAL).
 RP MEDLINE=75059123; PubMed=4803843;
 RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
 RT "The primary structure of a monoclonal IGM-immunoglobulin
 RT (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-
 RT type), subgroup H III. Architecture of the complete IGM-molecule.";
 RT

RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509 (1973).

[3]

RP REVISIONS (GAL).

RX MEDLINE=81066716; PubMed=6777162;

RA Mihaesco E., Barnikol-Watanabe S., Barnikol H.U., Mihaesco C., Hiltschmann N.,

RA "The primary structure of the constant part of mu-chain-disease protein BOT".

RT Eur. J. Biochem. 111:275-286 (1980).

RL [4]

RP SEQUENCE (WALDENSTROM'S OU), DISULFIDE BONDS, AND CARBOHYDRATES.

RX MEDLINE=74005511; PubMed=4742735;

RA Putnam F.W., Florent G., Paul C., Shinoda T., Shimizu A.,

RA "Complete amino acid sequence of the mu heavy chain of a human IgM immunoglobulin".

RT Science 182:287-291 (1973).

RL [5]

RP PARTIAL SEQUENCE FROM N.A.

RX MEDLINE=82059479; PubMed=6795593;

RA Rabbits T.H., Forster A., Milstein C.P.,

RA "Human immunoglobulin heavy chain genes: evolutionary comparisons of C mu, C delta and C gamma genes and associated switch sequences.";

RL Nucleic Acids Res. 9:4509-4524 (1981).

RL [6]

RP SEQUENCE OF 299-387 AND 438-454 FROM N.A.

RX MEDLINE=81077306; PubMed=6777778;

RA Dolby T.W., Devuono J., Croce C.M.,

RA "Cloning and partial nucleotide sequence of human immunoglobulin mu chain cDNA from B cells and mouse-human hybridomas.";

RT Proc. Natl. Acad. Sci. U.S.A. 77:6027-6031 (1980).

RL [7]

CC -!- MISCELLANEOUS: ALL 4 COMBINATIONS OF THE S/G & V/G POLYMORPHISMS AT POSITIONS 192 AND 216 HAVE BEEN OBSERVED IN HUMAN MU CHAINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; X17115; CAA34971.1; ALT SEQ.

DR EMBL; X57086; -; NOT_ANNOTATED_CDS.

DR PIR; A02162; MHU.

DR HSSP; P01857; 1FCL.

DR GENE; HGNC:5541; IGHM.

DR MIM; 147020; -.

DR GlycoSuiteDB; P01871; -.

DR InterPro; IPR003006; Ig MHC.

DR InterPro; IPR003597; Ig cl.

DR InterPro; IPR003600; Ig_like.

DR Pfam; PF00047; Ig; 4.

DR SMART; SM00410; Ig like; 1.

DR SMART; SM00407; IGcl; 3.

DR PROSITE; PS00290; IG MHC; 3.

KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.

FT NON_TER 1 1

FT DOMAIN 1 105

FT DOMAIN 106 218

FT DOMAIN 219 324

FT DOMAIN 325 454

FT DISULFID 14 14

FT DISULFID 28 88

FT DISULFID 135 198

FT DISULFID 215 215

FT DISULFID 245 304

FT DISULFID 292 292

FT CH1.

FT CH2.

FT CH3.

FT CH4.

INTERCHAIN (WITH A LIGHT CHAIN).

INTERCHAIN (WITH A HEAVY CHAIN).

INTERCHAIN (WITH A HEAVY CHAIN) IN ANOTHER OF THE 5 TETRAMERIC SUBUNITS OF THE MOLECULE).

FT DISULFID 352 414

FT DISULFID 453 453

FT CARBOHYD 46 46

FT CARBOHYD 210 210

N-LINKED (GLCNAC. . .).

FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .).

FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).

FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .).

FT VARIANT 192 192 /FTID=CAR_000219.

FT S -> G.

FT VARIANT 216 216 /FTID=VAR_003903.

FT V -> G.

FT /FTID=VAR_003904.

SQ SEQUENCE 454 AA; 49556 MW; 21EC72EADC56922E CRC64;

Query Match 26.9%; Score 499; DB 1; Length 454;

Best Local Similarity 31.5%; Pred. No: 1.5e-30;

Matches 106; Conservative 72; Mismatches 135; Indels 24; Gaps 9;

Qy 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSFPAKVHVTVLWVDOGAEN 69

Db 102 PLPVIAELPPKVSFV---VPRDGFPGNPKSKLICOATGFSRQIOVSWLRGKQVGS 158

Qy 70 LFPYTT-----RPKREGGQTFSLQSEVNITQGGWMSNTYTCVKHNGSIP-EDSSRRCS 123

Db 159 --GVTTDQVQAEAKESGPTTYKVTITLIKESDWLSQSMFTCRVDHRLTFQGNASMCV 216

Qy 124 DDEPRGVITYLIPSPDLIYENGTPKLTCLVLDSEENITVTVVRKKSIGASQSRST 183

Db 217 PDQDTAIRVFAIPPSFASIFLTSTKLTCLVTLDTLTYDVSVTISWTRQNGEAVKTHTNISE 276

Qy 184 KHHHATTISILPVDADKMEGEGYOCYRDVHPFKPIVRSITKLPKRL-APEVYMLP 242

Db 277 SHPNATFSVAGEASICEDDWSNGSRFTCTVTHTDLPSLQKQTSRPGKVALHRPDVILLP 336

Qy 243 PSPEETG--TTRTWTCLIRGFSEISVQWLPNNEEDHTGHTTTTRPKQKHGTDPSEFLY 300

Db 337 PAREQLNRESATITCLVTGFSADVFQWQRCQPLSPEKYVTSAPMPPEQAPGRYFAH 396

Qy 301 SRMLVNKSIWEKGNLVTCTRVVHEALPG---SRTLEKS 334

Db 397 SILTVSEEWNTGETYTCVVAHEALPNRVTERTVDSK 433

RESULT 5

MUCM RABIT STANDARD; PRT; 479 AA.

AC P04221;

DT 20-MAR-1987 (Rel. 04, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig mu chain C region membrane-bound form.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A. (A2 ALLOTYPED).

RX MEDLINE=84088930; PubMed=6418803;

RA Bernstein K.E., Alexander C.B., Reddy E.P., Mage R.G.;

RA "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of Vha2 allotype: comparisons with Vha1 and membrane mu sequences.";

RT J. Immunol. 132:490-495 (1984).

RL

CC -!- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM. THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-TERMINAL SEGMENTS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL; K01357; AAA31293.1; -.

DR PIR; A02165; MHRBM.

```

DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003597; Ig C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00410; Ig_Like; 2.
DR PROSITE; PS00290; IG MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON TER. 1
FT DOMAIN 1 106
FT DOMAIN 107 222
FT DOMAIN 223 327
FT DOMAIN 328 458
FT TRANSMEM 459 476
FT DISULFID 14 14
FT DISULFID 28 90
FT DISULFID 137 200
FT DISULFID 219 219
FT DISULFID 249 308
FT DISULFID 296 296
FT DISULFID 356 418
FT CARBOHYD 46 46
FT CARBOHYD 114 114
FT CARBOHYD 212 212
FT CARBOHYD 261 261
FT CARBOHYD 277 277
FT CARBOHYD 284 284
FT CARBOHYD 445 445
SQ SEQUENCE 479 AA; 52351 MW; 689C637A47B19FC CRC64;

Query Match 26.4%; Score 490; DB 1; Length 479;
Best Local Similarity 32.4%; Pred. No. 7.7e-30;
Matches 115; Conservative 62; Mismatches 154; Indels 24; Gaps 9;

Qy 5 HHHHTLSLPSGVPV-TIIPPTVKLFHSSCDPR-----GDAHSTIQLCLVSGFSPAKVHT 59
Db 94 HSNRNRLRVFPVDSLPVNSVF---IPRDSFGSGTRKSLICQATGFSKQISVS 150

Qy 60 WLVDGQAEEN---LFPYTRPKREGGQTFSLQSEVNTIQGOWMSNTYTCVHKNGSIFE 116
Db 151 WLRDQKVESGLTKPVEAEATKGAGPATFSLSSMLTITESDLWSQSLYTCRVDHRGIFPD 210

Qy 117 DS-----SRCSDEPRGVITYLIPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKK 173
Db 211 KNVSMSSBCTTPSPGIQVFPFIAPSFADTFLSKARLICLVTLDTTYGSLNLSWASHNGK 270

Qy 174 SIGASQRSTKHHTATTSITSLPVDADKWTIEGEGYQCRVDHPHPKPIVRSITKLPGR 233
Db 271 ALDTHMNTSHPNATFSAMGEASVCAEDWESGQFTCTVTHADLPPLPKHTISK--SRE 328

Qy 234 LA---PEVYMLPPSPPEE--TGTTRTVTLIRGFYPSSEISVQWLPNNEEDHTGHTTTRPQ 288
Db 329 VAKIPPAVYLPAPAREQVLRESATVCLVKGFSADVFQVQWQGRGQLPSDKKVTSGAPA 388

Qy 289 KDHGTDPSFFLYSLRMVNKSIWEKGNLVTCTRVHEALP---GSRLEKSLHYSAG 340
Db 389 PEQAPGLYFHTSLITVTEEDWNSGETFTCVVGHGHEALPHMVTERTVDSKSTEGEVG 443

RESULT 6
MUC_RABIT
ID MUC_RABIT STANDARD; PRT; 458 AA.
AC P03988;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig mu chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_taxid=9986;
RN [1]

```

```

RP SEQUENCE FROM N.A. (A2 ALLOTYPED).
RX MEDLINE=8408930; PubMed=6418803;
RA Bernstein K.B.; Alexander C.B.; Reddy E.P.; Mage R.G.;
RT "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain
of VhA2 allotype: comparisons with VhA1 and membrane mu sequences.";
RL J. Immunol. 132:490-495(1984).
CC -1- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES
CC SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM.
CC THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-
CC TERMINAL SEGMENTS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K01357; -; NOT_ANNOTATED_CDS.
CC PIR; A02164; MHRB.
CC HSSP; P01842; 7FAB.
CC InterPro; IPR003006; Ig MHC.
CC InterPro; IPR003597; Ig_C1.
CC InterPro; IPR003600; Ig_Like.
CC Pfam; PF00047; Ig; 4.
CC SMART; SM00410; Ig_Like; 2.
CC SMART; SM00407; IGEL; 2.
CC PROSITE; PS00290; IG MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON TER. 1
FT DOMAIN 1 106
FT DOMAIN 107 222
FT DOMAIN 223 327
FT DOMAIN 328 458
FT DISULFID 14 14
FT DISULFID 28 90
FT DISULFID 137 200
FT DISULFID 219 219
FT DISULFID 249 308
FT DISULFID 296 296
FT DISULFID 356 418
FT CARBOHYD 46 46
FT CARBOHYD 114 114
FT CARBOHYD 212 212
FT CARBOHYD 261 261
FT CARBOHYD 277 277
FT CARBOHYD 284 284
FT CARBOHYD 445 445
SQ SEQUENCE 458 AA; 49897 MW; F338D6A3B28B4864 CRC64;

Query Match 26.3%; Score 488; DB 1; Length 458;
Best Local Similarity 32.7%; Pred. No. 1e-29;
Matches 114; Conservative 62; Mismatches 149; Indels 24; Gaps 9;

Qy 5 HHHHTLSLPSGVPV-TIIPPTVKLFHSSCDPR-----GDAHSTIQLCLVSGFSPAKVHT 59
Db 94 HSNRNRLRVFPVDSLPVNSVF---IPRDSFGSGTRKSLICQATGFSKQISVS 150

Qy 60 WLVDGQAEEN---LFPYTRPKREGGQTFSLQSEVNTIQGOWMSNTYTCVHKNGSIFE 116
Db 151 WLRDQKVESGLTKPVEAEATKGAGPATFSLSSMLTITESDLWSQSLYTCRVDHRGIFPD 210

Qy 117 DS-----SRCSDEPRGVITYLIPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKK 173
Db 211 KNVSMSSBCTTPSPGIQVFPFIAPSFADTFLSKARLICLVTLDTTYGSLNLSWASHNGK 270

Qy 174 SIGASQRSTKHHTATTSITSLPVDADKWTIEGEGYQCRVDHPHPKPIVRSITKLPGR 233
Db 271 ALDTHMNTSHPNATFSAMGEASVCAEDWESGQFTCTVTHADLPPLPKHTISK--SRE 328

Qy 234 LA---PEVYMLPPSPPEE--TGTTRTVTLIRGFYPSSEISVQWLPNNEEDHTGHTTTRPQ 288

```

```
Db 329 VAKHPAVVVLPPAREQLVLRASATVTCVLKGFSPADVFQVQQRGQPLSSDKVYTSAPA 388
QY 289 KDHGTDPSFFLYSRMLVYKSIWEKGNLVTCTRVVHEALP---GSRITLKS 334
Db 389 PEQAPGLYFTHSTLTVTEEDWNSGETFTCVVGHEALPHMVTERTVDKS 437

RESULT 7
MUCB_HUMAN
ID MUCB_HUMAN STANDARD; PRT; 391 AA.
AC P04220;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig mu heavy chain disease protein (BOT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=84184186; PubMed=6425189;
RA Barnikol-Watanabe S., Mihaesco E., Mihaesco C., Barnikol H.U.,
RA Hilschmann N.;
RT "The primary structure of mu-chain-disease protein BOT. Peculiar
RT amino-acid sequence of the N-terminal 42 positions.";
RL Hoppe-Seyler's Z. Physiol. Chem. 365:105-118 (1984).
CC -!- MISCELLANEOUS: THIS PROTEIN HAS NO V REGION HOMOMOLOGY OR CHI
CC REGION.
DR PIR; A02163; MHUJBT.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGG1; 2.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 42
FT PRE-C-PART (NO V REGION HOMOMOLOGY).
FT DOMAIN 43 155
FT CH2.
FT DOMAIN 156 261
FT CH3.
FT DOMAIN 262 391
FT CH4.
SQ SEQUENCE 391 AA; 43057 MW; 9100843AF0CF021A CRC64;

Query Match 26.1%; Score 484.5; DB 1; Length 391;
Best Local Similarity 31.2%; Pred. NO. 1.6e-29;
Matches 105; Conservative 72; Mismatches 137; Indels 23; Gaps 10;

QY 15 SGPVTI-IPPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHVTVLVDGQEAEN 69
Db 40 SQPVIAELPPKVSVP---VPRDGGFGNPRKS-KLICQATGSPKQIEVSWLRGKQVGS 95
QY 70 LPFYTT-----RPKREGQTSLQSEVNITQGMSSNTYTCVKNHGSIF-EDSSRRCS 123
Db 96 --GVTTDEVEAEKESGTYTKVYTLTKESDWLGQSMFTCRVDHRLGLTQQNASSCG 153
QY 124 DDEPGVITYLIPPSPLDLYNGTKLCLVLDSEENITVTVVREKKSIGSASQST 183
Db 154 PQDQTAIRVFAIPPSFASIFLTSTKLTCLVTLDTTYSVTISWTRQDGEAVKTHNISE 213
QY 184 KHHHATTITSILPVDADKWIEGEGYQCRVDPHPKPIVRSITKLPCKRL-APEVYMLP 242
Db 214 SHPNATFSAVGEASICEEDWDSGERFTCTVTHDLPSPKQITSRPKGVALLHRPDVLLP 273
QY 243 PSPEETG--TTRTVTCLRGVPSISVQWLPNNDEETHGHTTTRPQKHGTDPSFFLY 300
Db 274 PAREQLNLRASATITCLVTGSPADVFQVQQRGQPLSPKXVTSAPWPEQAPGRYFAH 333
QY 301 SRMLVYKSIWEKGNLVTCTRVVHEALP---SRTLEKS 334
Db 334 SILTVSBEWNTGETYTCVVAHEALPNRVTERTVDKS 370
```

```
RESULT 8
GC2_HUMAN
ID GC2_HUMAN STANDARD; PRT; 326 AA.
AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-2 chain C region.
GN IGHG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 2-326 FROM N.A.
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
RT heavy chain constant region genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988 (1982).
RN [2]
RP SEQUENCE OF 88-115 FROM N.A.
RX TISSUE=Fetal liver;
RP MEDLINE=83001943; PubMed=6811139;
RA Takahashi N., Ueda S., Obata M., Nikaide T., Nakai S., Honjo T.;
RT "Structure of human immunoglobulin gamma genes: implications for
RT evolution of a gene family.";
RL Cell 29:671-679 (1982).
RN [3]
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.
RX TISSUE=Fetal liver;
RP MEDLINE=84235992; PubMed=6329676;
RA Krawinkel U., Rabbitts T.H.;
RT "Comparison of the hinge-coding segments in human immunoglobulin gamma
RT heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
RT genes.";
RL EMBO J. 1:403-407 (1982).
RN [4]
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX MEDLINE=81007873; PubMed=6774012;
RA Wang A.-C., Tung E., Fudenberg H.H.;
RT "The primary structure of a human IgG2 heavy chain: genetic,
RT evolutionary, and functional implications.";
RL J. Immunol. 125:1048-1054 (1980).
RN [5]
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
RX MEDLINE=80001357; PubMed=113060;
RA Connell G.E., Parr D.M., Hofmann T.;
RT "The amino acid sequences of the three heavy chain constant region
RT domains of a human IgG2 myeloma protein.";
RL Can. J. Biochem. 57:758-767 (1979).
RN [6]
RP SEQUENCE OF 238-275 (ZIE).
RX MEDLINE=80114419; PubMed=118920;
RA Hofmann T., Parr D.M.;
RT "A note of the amino acid sequence of residues 381-391 of human
RT immunoglobulin gamma chains.";
RL Mol. Immunol. 16:923-925 (1979).
RN [7]
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
RA Hofmann T., Parr D.M.;
RL Submitted (MAR-1980) to the PIR data bank.
RN [8]
RP SEQUENCE OF 1-121 (DOT).
RX MEDLINE=95255298; PubMed=7737190;
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-flavin monoclonal
RT immunoglobulins.";
RL Eur. J. Biochem. 228:886-893 (1995).
RN [9]
RP DISULFIDE BONDS.
```


RX MEDLINE=72033500; PubMed=4940472;
 RA Milstein C., Frangione B.;
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
 RL Biochem. J. 121:217-225(1971).
 RN [10]
 RP DISULPHIDE BONDS.
 RA MEDLINE=69064124; PubMed=5782707;
 RA Frangione B., Milstein C., Pink J.R.L.;
 RT "Structural studies of immunoglobulin G.";
 RL Nature 221:145-148(1969).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J00230; AAB59393.1; -;
 DR PIR; A02148; G2HU.
 DR HSPP; P01857; 1FC1.
 DR Genew; HGNC:5526; IGHG2.
 DR MIM; 147110;
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_cl.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00410; Ig like; 1.
 DR SMART; SM00407; IGcl; 2.
 DR PROSITE; PS00290; IG_MHC; 2.
 DR Immunoglobulin domain; Immunoglobulin C region.
 KW NON TER 1 98
 FT DOMAIN 1 98
 FT CH1.
 FT HINGE.
 FT DOMAIN 99 110
 FT DOMAIN 111 219
 FT DOMAIN 220 326
 FT DISULFID 14 14
 FT DISULFID 27 83
 FT DISULFID 102 102
 FT DISULFID 103 103
 FT DISULFID 106 106
 FT DISULFID 109 109
 FT DISULFID 140 200
 FT DISULFID 246 304
 FT SITE 156 156
 FT MOD RES 326 326
 FT VARIANT 60 60
 FT CONFLICT 109 109
 FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;
 Query Match 25.2%; Score 468; DB 1; Length 326;
 Best Local Similarity 34.4%; Pred. No. 2.2e-28;
 Matches 111; Conservative 66; Mismatches 116; Indels 30; Gaps 12;
 Qy 35 RGDASTIQLCLVSGSPAKVHTW----LVDQAEANLFPYTRPKREGGQTFSLQSE 90
 Db 16 RSTSESTAALGCLVKDYFPEPTVSWNSGALTSG---VHTPAVL-----QSSGLYSLSSV 68
 Qy 91 VNITQGQMSNTYTVCHVKH---NGSIPEDSSRCSDDEP-----RGVITYLIPPSPL 140
 Db 69 VTPSSNF-GRQYTCNVDDHPSNTKVDKTVKCCVCPAPPVAGPSVFLFPKPK 127
 Qy 141 D-LYENGTPKLTCLVLDLSESE-NITVTWVRKKSIGSASQRTKHHATTSITSLPV 198
 Db 128 DTLMSIRTPETVTCVVDVSHSDPEVQFNWYVDGVHNAKTKPREEQNSFRVVSVLTV 187
 Qy 199 DAKWIEGEGYQCRVDHDFPKPIVRSITKLPGRKLAPEVYMLPPSPPE-TGTRTITVCL 257
 Db 188 VHQDLNGKEYCKVSNKGLPAPEKTIISKTKGQPREQVTVLPSPREEMTKNQVSLTCL 247
 Qy 258 IRGYPFSEISVQWMLPNNEEDHTGHHHTTRTPQKHGTDPSFFLYSRMLVNKSIWEKGNLVT 317

Db 248 VKGFPSDIAVEWESNQPEN--NYKTPPMLD--SDGSFFLYSKLVTKSRWQGNVFS 303
 Qy 318 CRVVEALPGSRITKLSLHYSAG 340
 Db 304 CSVWEALHNHYT-QKSLSLSPG 325
 RESULT 9
 MUC_CANFA
 ID_MUC_CANFA STANDARD; PRT; 450 AA.
 AC P01874; 1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig mu chain C region.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE OF 1-177 (MOO).
 RX MEDLINE=80077682; PubMed=117299;
 RA McCumber L.J., Capra J.D.;
 RT "The complete amino-acid sequence of a canine mu chain.";
 RL Mol. Immunol. 16:565-570(1979).
 RN [2]
 RP SEQUENCE OF 178-450 (MOO).
 RX MEDLINE=78180587; PubMed=653360;
 RA Wasserman R.L., Capra J.D.;
 RT "Amino acid sequence of the Fc region of a canine immunoglobulin M:
 RT interspecies homology for the IgM class.";
 RL Science 200:1159-1161(1978).
 DR PIR; A02169; MHDG.
 DR HSPP; P01857; 1FC1.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_cl.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig; 4.
 DR SMART; SM00410; Ig like; 2.
 DR SMART; SM00407; IGcl; 2.
 DR PROSITE; PS00290; IG_MHC; 3.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON TER 1
 FT SEQUENCE 450 AA; 48895 MW; 9D460DA9D1012F5D CRC64;
 Query Match 25.0%; Score 464.5; DB 1; Length 450;
 Best Local Similarity 31.8%; Pred. No. 6e-28;
 Matches 109; Conservative 69; Mismatches 136; Indels 29; Gaps 12;
 Qy 10 LSLPE--SGPVTIPTPVTKLFHSSCDPRGDASHSTIQLCLVSGSPAKVHTW-LVDQGE 66
 Db 106 LTPPEVSG--FIPPRDAFG--BPKS-----QLICQASGSPQV---WSLRGKQ 151
 Qy 67 AENLFPYTT-----RPKREGQTFSLQSEVNITQGQMSNTYTVCHVKHNSIF-EDSSR 120
 Db 152 IES--GVTTNEVZAZAKZSGPTTVKVTMLTIQEDAMLSQSVFTCKVEHRLTQQNASS 209
 Qy 121 RCDSDPERGVITVLIIPSPDLVYENGTPKLTCLVLDLSESENTVTWVRKKSIGSASQ 180
 Db 210 MCTSDQVGVISITFIPPSFASIFNTKSAKLSCLVTDLATYDSVTISTRENGALKHTTN 269
 Qy 181 RSTKHHATTISITLVPDAKDMTEGEGYQCRVDHDFPKPIVRSITKLPGRKL-APEVY 239
 Db 270 ISSHPNGTFSAGEATVCVEWESGEQFTCTVHTDLPVLKOTISRPKGVAVHMSVY 329
 Qy 240 MLPPSPPEET--TTRTVTLIRGFYPSSEISVQWMLPNNEEDHTGHHHTTRTPQKHGTDPSF 297
 Db 330 VLPPSRQLDRESATLSCLTVGSPDPVQVQKQPPVPPDSYVTSAPMPEQAPGLY 389
 Qy 298 FLYSRMLVNKSIWEKGNLVTCTVVEALPGSRITKLSLHYSAG 340
 Db 390 FAHSILTVEEENAGETVTCVVAHESLP-NRVTESVDKSTG 431

```
RESULT 10
MUC_MOUSE
ID MUC_MOUSE STANDARD; PRT; 455 AA.
AC P01872;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig mu chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81076590; PubMed=6255422;
RA Kawakami T., Takahashi N., Honjo T.;
RT "Complete nucleotide sequence of mouse immunoglobulin mu gene and
RT comparison with other immunoglobulin heavy chain genes.";
RL Nucleic Acids Res. 8:3933-3945(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82051295; PubMed=6795090;
RA Goldberg G.I., Vanin E.F., Zrolka A.M., Blattner F.R.;
RT "Sequence of the gene for the constant region of the mu chain of
RT Balb/c mouse immunoglobulin.";
RL Gene 15:33-42(1981).
RN [3]
RP SEQUENCE FROM N.A. (MYELOMA TEPCL183).
RX MEDLINE=81165562; PubMed=6260591;
RA Auftray C., Rougeon F.;
RT "Nucleotide sequence of a cloned cDNA corresponding to secreted mu
RT chain of mouse immunoglobulin.";
RL Gene 12:77-86(1980).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN MOPC 104E).
RX MEDLINE=79223904; PubMed=111247;
RA Kehry M.R., Sibley C.H., Fuhrman J.S., Schilling J.W., Hood L.E.;
RT "Amino acid sequence of a mouse immunoglobulin mu chain.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2932-2936(1979).
RN [5]
RP REVISION (MOPC 104E).
RX MEDLINE=83075344; PubMed=6816276;
RA Kehry M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
RA Hood L.E.;
RT "Complete amino acid sequence of a mouse mu chain: homology among
RT heavy chain constant region domains.";
RL Biochemistry 21:5415-5424(1982).
CC -!- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES
CC SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM.
CC THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-
CC TERMINAL SEGMENTS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: V00827; -; NOT_ANNOTATED_CDS.
CC PIR: A02166; MHMS.
CC HSP: P01857; IFC1.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003597; Ig_c1.
CC InterPro: IPR003600; Ig_like.
CC Pfam: PF00047; Ig_4.
CC SMART: SM00410; Ig_like; 2.
CC SMART: SM00407; Ig_c1; 2.
CC PROSITE: PS00290; Ig_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
```

```
Alternative splicing.
KW NON TER 1
FT DOMAIN 1 105 CH1.
FT DOMAIN 106 217 CH2.
FT DOMAIN 218 324 CH3.
FT DOMAIN 325 436 CH4.
FT DOMAIN 437 455
FT DISULFID 14 14
FT DISULFID 28 89
FT DISULFID 136 199
FT DISULFID 216 216
FT DISULFID 246 305
FT DISULFID 293 293
FT DISULFID 353 415
FT DISULFID 454 454
FT CARBOHYD 46 46
FT CARBOHYD 211 211
FT CARBOHYD 243 243
FT CARBOHYD 281 281
FT CARBOHYD 442 442
FT VARIANT 78 78
FT VARIANT 101 101
FT VARIANT 226 226
FT VARIANT 258 258
FT VARIANT 258 258
FT VARIANT 368 368
SQ SEQUENCE 455 AA; 50101 MW; 4CB57CB602F9B51 CRC64;

Query Match 24.5%; Score 455.5; DB 1; Length 455;
Best Local Similarity 32.1%; Pred. No. 2.9e-27;
Matches 105; Conservative 63; Mismatches 140; Indels 19; Gaps 8;

Qy 23 PTVKLFHSSCDPR----GDAHSTIQLCLVSGSPAKVHVTVLVDGQEAENLF---PVT 75
Db 112 PNVNVF---VPPRGFGSPAPRKSLICEATNFTPKPTVSWLKGDLVESGFTDPVTI 168

Qy 76 RPKREGGQTFSLQSEVNITQGMSSNTYCHVKHNGSIF-EDSSRRCSDDEPRGVITYL 134
Db 169 ENKGSTPTQYKIVSTLTISEIDMLNLYTCRVDRHGLTFLKNVSSSTCAASPSDILFT 228

Qy 135 IPSPLDLYENGPKLTCLVLDLSEENITVTVRERKKSIGSASQSTKHHTTSITS 194
Db 229 IPPSFADIFLKSANLTCLVSNLATYETLNIWSASQSGEPLTKIKIMESHGPNFTSAGK 288

Qy 195 ILPVDKDWIEGEGYQCRVDHPHPKPIVRSITKLPCK--RLAPEVTMLPSPPEETG--T 250
Db 289 VASVCVEDMNNRKEFVCTVTHRDLPSQKFKIS-K-PNEVHKHPAPVYLLPAREQLNRE 347

Qy 251 TRVTCLIRGFPSEISVQWLPNNEEDHTGHTTTRPKDHGTDPSFFLYSRMLVNKSIW 310
Db 348 SATVTCVLKGFSPADISVQMLQGLLPQEKYVTSAPMPEPGAPGFYFTHSILTVTEEW 407

Qy 311 EKNLVTCTRVVHRLP---GSRILEKS 334
Db 408 NSGETYTCVVGHEALPHLVTERTVDKS 434

RESULT 11
MUC_MOUSE
ID MUC_MOUSE STANDARD; PRT; 476 AA.
AC P01873;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig mu chain C region membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 433-476 FROM N.A.
RX MEDLINE=80222874; PubMed=6771020;
RA Early P., Rogers J., Davis M., Calame K., Bond M., Wall R.,
```



```
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811BF208E7A CRC64;
Query Match 24.3%; Score 451.5; DB 1; Length 327;
Best Local Similarity 34.3%; Pred. No. 3.9e-27;
Matches 111; Conservative 62; Mismatches 120; Indels 31; Gaps 11;
QY 35 RGDHSTTQLLCLVSGFSPAKVHVTVW---LVDGOEAEHLFPYTRPRKGGQTFSLQSE 90
DB 16 RSTSESTAALGCLVKDYPEPVTWVSNWNGALTSG---VHTPAVL---QSSGLYSLSV 68
QY 91 VNITQGMWSSNTYCHVKNHSGIPEDSSRCSDDEP-----RGVITLIPSP 139
DB 69 VTFPSSS-LGTYTCVNDHKNPSNTKVDKRVESKYGPCCPAPCEFLGSPSFLFPKP 127
QY 140 LD-LYENCTPKLTCVLVLESEE-NITVTVWRERKKSIGSASQSTKHHTTSTSLP 197
DB 128 KDTLMISRTPEVTCVVDVDSQEDPEVQFNWYDGVVEVHNATKPREQFNSTYRVSVLT 187
QY 198 VDADWIEGEGYQCRVDHPHPKPIVRISITKLPGRKLAPEVYMLPPSPPEE-TGTRTVTC 256
DB 188 VLHQDLNGKEYCKVSNKGLPSSTIEKTSKAKGPREPQVYTLPPSQEEMTKQVSLTC 247
QY 257 LIRGYPSEISVQMLPNNEEDHTHTTRPKQKHGTDPSFLYSRMLVKNKSIWEKNLV 316
DB 248 LVKGYPSDIAVEWESNGQPN--NYKTPPVLDD--SDGSFSLYSLRTVDSRMQEGNVF 303
QY 317 TCRVVHEALPGSRTLEKSLHYSAG 340
DB 304 SCSVWHEALHNYT--QKSLSLSLG 326

RESULT 13
ID MUC_SUNMU STANDARD; PRT; 457 AA.
AC P20768;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig mu chain C region.
OS Suncus murinus (House shrew) (Musk shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Soricidae; Crocidurinae; Suncus.
OX NCBI_TaxID=9378;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver.
RX MEDLINE=89232144; PubMed=2497033;
RA Ishiguro H., Ichihara Y., Namikawa T., Nagatsu T., Kurosawa Y.;
RT "Nucleotide sequence of Suncus murinus immunoglobulin mu gene and
RL comparison with mouse and human mu genes.";
FEBS Lett. 247:317-322(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13920; CAA32113.1; ALT_INIT.
DR PIR; S03961; S03961.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGCL; 3.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 105
FT DOMAIN 1 105 CH1.
```

```
FT DOMAIN 106 220 CH2.
FT DOMAIN 221 326 CH3.
FT DOMAIN 327 457 CH4.
FT DISULFID 13 13 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 27 89 BY SIMILARITY.
FT DISULFID 136 200 BY SIMILARITY.
FT DISULFID 217 217 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 248 307 BY SIMILARITY.
FT DISULFID 295 295 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 355 417 BY SIMILARITY.
FT DISULFID 456 456 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 457 AA; 50074 MW; 56C8C086DA4462E9 CRC64;
Query Match 24.2%; Score 449.5; DB 1; Length 457;
Best Local Similarity 29.9%; Pred. No. 8.3e-27;
Matches 100; Conservative 76; Mismatches 138; Indels 21; Gaps 8;
QY 21 IPPTVKLFHSSCDPR----GDAHSTIQLCLVSGFSPAKVHVTVLVDGOEAEHLFPYTT- 75
DB 110 LPPNVISF--VPPRNSFGNHPRTSOLICQASGSPRTIVMSLQGEVPQPSLVSTSA 166
QY 76 ---RPREGGQTSLOSEVNITQGMWSSNTYCHVKNHSGIPED--SSRCSDEPRGV 130
DB 167 VEAPKSGGPTTFRVISRLTITENWLSQREFTCQALHKLGLTFOKNVSSVCMGDDTSTGI 226
QY 131 ITVLPPSLDLYENGTPKLTCLVLDLESENITVTVWRERKKSIGSASQSTKHHTT 190
DB 227 SVFLPPTFANIFLTQSAQLTCLVTGLATYDSLDSISWRQNGEALQTHVISEHPNSTF 286
QY 191 SITSLPVDADWIEGEGYQCRVDHPHPKPIVRISITKLPGRKLA---PEVYMLPPSPPEE 247
DB 287 TAKGHASVCREEWESEGEKCTCTVQHSPLSPKLSLR--PKDVANDPPSVFVLPQAQEQ 344
QY 248 --TGTTTCTVCLIRGFYPSISVQMLPNNEEDHTHTTRPKQKHGTDPSFLYSRMLV 305
DB 345 LKUREASITCLVKDFPDPVQVQHQHQFVDPKHYVTNSPTPEPONPGFLYVHSILTV 404
QY 306 NKSIWEKGNLVTCTRVVHEALPGSRTLEKSLHYSAG 340
DB 405 SEKDWSGESFSCVVGHEALPLSVT-EKAVDKTSG 438
RESULT 14
ID GC1_HUMAN STANDARD; PRT; 330 AA.
AC P01857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-1 chain C region.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP SEQUENCE OF 1-135 (MYELOMA PROTEIN BU).
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA Waxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
```

[illegible]

Db 58 --QSSGLYSLSSVWTPVSS--LGTQYICNVNHKPSNTKVDKKBPKSCDKHTTCCPPCPA 114
Qy 124 DDEPRGVITYLPPSPFLD-LYENGTPKLTCLVLDSEBE-NITVTVWRERKKSIGSASOR 181
Db 115 PELGCGPSVFLPPPKTKTLMISRPETVCVVDVSHEDPEVKFNWYVDGVEVHNATKP 174
Qy 182 STKHHHTATTSITSLPVDADAKWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYML 241
Db 175 REEQYNSTRVSVVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 234
Qy 242 PPSPEE-TGTRTVTCLIRGYPSPISVQWLPNNEEDHTGHTTTPRQKDHGTDBSFPLY 300
Db 235 PPSRDLTKNQVSLTKLVKGFYPSDAVWESNGQPEN--NYKTPPVLD--SDGSFFLY 290
Qy 301 SRMLYNKSIWEKGNLVTQVHVHEALPGRSTLEKSLHYSAG 340
Db 291 SKLTVDKSRWQGNVFSVWHEALHNHYT-QKSLSLSPG 329

RESULT 15

MUC_MESAU
ID MUC_MESAU STANDARD; PRT; 454 AA.
AC P06337;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig mu chain C region.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85297761; PubMed=2994005;
RA McGuire K.B., Duncan W.R., Tucker P.W.;
RT "Phylogenetic conservation of immunoglobulin heavy chains: direct
comparison of hamster and mouse Cmu genes.";
RL Nucleic Acids Res. 13:5611-5628(1985)
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02804; CAA26574.1; -.
DR PIR; A02168; MHY.
DR HSP; P01854; 1IGF.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00410; Ig_Like; 2.
DR SMART; SM00407; IG_C1; 2.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 105 CH1.
FT DOMAIN 106 218 CH2.
FT DOMAIN 219 324 CH3.
FT DOMAIN 325 454 CH4.
FT DISULFID 13 13 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 27 88 BY SIMILARITY.
FT DISULFID 135 198 BY SIMILARITY.
FT DISULFID 215 215 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 245 304 BY SIMILARITY.
FT DISULFID 292 292 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 352 414 BY SIMILARITY.
FT DISULFID 453 453 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 454 AA; 50196 MW; 4DA2134612BE1469 CRC64;
Query Match 23.3%; Score 433.5; DB 1; Length 454;
Best Local Similarity 30.7%; Pred. No. 1.3e-25;
Matches 106; Conservative 68; Mismatches 154; Indels 17; Gaps 9;
Qy 5 HHHHT--LSLPSSGPTIIPPTVKLFHSSCDP-RGDAHSTIQLLCLVSGFSPAKVHVTV 60
Db 91 HHGNTNKDLAVP-IPVTENNPNVSVFVPSRDASGAPAKRSRLFCEASNFSPKQITVSW 149
Qy 61 LVGQEAENLF---PYTRPKREGGQTFSLQSEVNIITOGOMMSNTVTVHVKHNGSIF-E 116
Db 150 LRDGKPVKSGFTTPEVTPEDRSGSPRTYKVIISTLTITESTDNLNLSVYTCRVDRHGLTFWK 209
Qy 117 DSSRRCSDDPRGVITYLIPSPDLPLYENGTPKLTCLVLDSEENITVTVWRERKKSIG 176
Db 210 NVSTCAASPTDIQAFPIPPSFVGFINKSATLUTCLVTNLATYDTLNLSSRSGBPLE 269
Qy 177 SASORSTKHHHTSITSLPVDADAKWIEGEGYQCRVDHPHPKPIVRSITKLPK--RL 234
Db 270 TKTKLTESHNGTFSALGEANVCVEDWDSGKEFVCTVTHRDLPSPQKKFISK-PRMNKT 328
Qy 235 APEVYMLPPSPPEE--TGTRTVTCLIRGYPSEISVQWLPNNEEDHTGHTTTPRQKDHG 292
Db 329 PPAVYQQPLAREQLILRESATVTCLVKGFSPADIFVQWLQKQGLSQDKYVTSAPMREPQ 388
Qy 293 TDPSFFLYSRMLYNKSIWEKGNLVTQVHVHEALP---GSRLEKS 334
Db 389 APHLYFTHSVLTVEEWNSSGETYTCVVGHEALPHMVTERTVDRS 433

Search completed: July 9, 2003, 13:40:01

Job time : 12 secs

Query Match	46.9%;	Score	871.5;	DB	2;	Length	548;
Best Local Similarity	53.9%;	Pred.	No. 9e-57;				
Matches	178;	Conservative	45;	Mismatches	92;	Indels	15;
Gaps							
Qy	12	LPESGPTTIIPTVKLFHSSCDPRGD--AHSITQLLCLVSGFSPAKVHVHTWLVGDGEAEN					
Db	229	VFEVSVFIPPKP-----DVLTRSTQIYCFIYGHILNDVSVWLMDDREITD					
Qy	70	LPFYTPRRREGQTFSLQSENIITQGMSSNTTCHVKHNGSIFEDSSRRCCSDDEPRG					
Db	279	TLAQTVLKEE--GKLASTCSKLNIITEQQMSESTFTCKVTSQGVLYLAHTRPCDPHEPRG					
Qy	130	VITYLIPSPDLYNGTPKLTCLVLDLSEENITVTVVRERKKSIGASQSRSTKH--HA					
Db	338	VITYLIPSPDLQYONGAPKLTCLVVDLESEKNVNVTWNQEKTSV--SASQWYTKHHNNA					
Qy	189	TTTSITSLPVDADKMTIEGEGYOCRDVHDHPKPIVRSITKLPGKRLAPVYMLPPSPERT					
Db	397	TTTSITSLPVAKDMIEGYQCIVDHPDPKPIVRSITKTPQGSAPENVYVFPPEEES					
Qy	249	GTTRVTCLIRGFYFSEISVQWLPNNEDHTGHHTTRTPQKDGHGTDGPFFLYSRMLVNKS					
Db	457	EDKRTLTCLIQNFPPEDISVQWLDGKULISNSQHSSTTTPKNSGNSRGFFIFISRLVAKT					
Qy	309	IWEKGNLVTCRVVHEALPGSRFTLEKSLHYS	338				
Db	517	LWTORKFTCOVTHEALOKPRLEKTI	546				

```

Db      517 LMTQRKQFTCVIHEALQPKRLEKTISTS 546

RESULT 4
EHMSS
IG epsilon chain C region (version 2) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1999
C/Accession: A02145
R:Ishida, N.; Ueda, S.; Hayaeshida, H.; Miyata, T.; Honjo, T.
EMBO J. 1, 1117-1123, 1982
A:/Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene:
A:/Reference number: A90966; MUID:84236092; PMID:6329728
A:/Accession: A02145
A:/Molecule type: DNA
A:/Residues: 1-423 <ISH>
A:/Note: the sequence was determined from the germline gene
C:/Genetics:
A:/Introns: 91/1; 199/1; 307/1
C:/Complex: An immunoglobulin heterotetramer subunit consists of two identical chains disulfide bonds. In some cases, such as IgA and IgM, the subunits associate superfamily: Immunoglobulin C region; immunoglobulin homology
C;/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin F;16-77/Domain: immunoglobulin homology <IMW1>
F;115-183/Domain: immunoglobulin homology <IMW2>
F;220-288/Domain: immunoglobulin homology <IMW3>
F;325-396/Domain: immunoglobulin homology <IMW4>
F;23-75,122-181,227-286,332-394/Disulfide bonds: #status predicted
F;43,84,167,239,262,417/Binding site: carbohydrate (Asn) #status
Query Match: 46.4%; Score 862.5; DB 1; Length 423;
Best Local Similarity 51.8%; Pred.No.3e-56;
Matches 183; Conservative 46; Mismatches 103; Indels 21; Gaps
QY      2 FHHHHTLSPESGVTITPTVKLFHS-----SCDPRGDASTIQALLCLVSGF
          |         :             |       |||||               :
          FCCHTHPFSNESRTILVRPVT----HSLSPWSVIHRCPDNA-FHSTIQLYCFYGH
Db      73
QY      52 SPAAKVHVTVLDQGAEANFLPYTPTRKREGGGTFSLQSEVNITQQGMSSNTYTTC HVGNK
          |         :             |       |||||               :

```


Db 128 ILNDVSVWLMDDREITDTLAQTVLKKEE-GKLASTCSKLNITEQQWMSBSTFTCRVTSQ 186
QY 112 GSIFEDSRRCSDDPRGVITTYLIIPSPDLIYENGTPKLTCLVLDLSEENITVWVRER 171
Db 167 GVDYLHTRRCPDHEPRGAIYYLIIPSPDLIYQNGAPKLTCLVDLSEKKNVNTWQEK 246
QY 172 KKSIGSASORSTKH-HATTSITSILPVDADKWIEGEGYQCRVDPHPKPIVRSITKLP 230
Db 247 KTSV-SASQWTKHNNATTSTSLPVVAKDWIEGEGYQCVDPDPKPIVRSIT-LP 304
QY 231 --GKELAEVVMPLPSPETGTRVTCILRGFPSEISVOMLNNNEEDHTGHHHTRPQ 288
Db 305 QVSRQSAEVYVFPPEESEDKRTLTCLIQNFPPEDISVOMLGDGKLISQHSHTTPL 364
QY 289 KDHGTDPEFFLYSRMLVKNKSWEKGNLVTCRVVHEALPGSRITLKSILHYSAGN 341
Db 365 KNSGNSQGFIFSRLEVAKTLTWQRKQFTQCQVIHEALQPKRLEKTIISTSLGN 417
RESULT 5
EHU
Ig epsilon chain C region - human
C/Species: Homo sapiens (man)
C/Date: 31-Mar-1981 #sequence_revision 13-Jun-1993 #text_change 16-Jul-1999
C/Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116; C46
R/Flanagan, J.G.; Rabbitts, T.H.
EMBO J. 1, 655-660, 1982
A/Title: The sequence of a human immunoglobulin epsilon heavy chain constant region gene
A/Reference number: A22771; MUID:84236029; PMID:6234164
A/Accession: A22771
A/Molecule type: DNA
A/Residues: 1-428 <FLA>
A/Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035
R/Ueda, S.; Nakai, S.; Nishida, Y.; Hiseajima, H.; Honjo, T.
EMBO J. 1, 1539-1544, 1982
A/Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudog
A/Reference number: A23195; MUID:84207910; PMID:6327276
A/Accession: A23195
A/Molecule type: DNA
A/Residues: 2-428 <UED>
A/Cross-references: GB:J00222; NID:g184755
R/Zhang, K.; Saxon, A.; Max, E.E.
J. Exp. Med. 176, 233-243, 1992
A/Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing
A/Reference number: PH1214; MUID:92308839; PMID:1613458
A/Accession: PH1214
A/Molecule type: DNA
A/Residues: 320-428 <ZHA>
A/Cross-references: EMBL:X63693; GB:S38668; NID:g32987
R/Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasada, R.; Kikuchi, M.; Sugii
Nucleic Acids Res. 11, 719-726, 1983
A/Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon cha
A/Reference number: A93491; MUID:83168897; PMID:6300763
A/Accession: A93491
A/Molecule type: mRNA
A/Residues: 1-428 <SEN>
A/Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035
R/Max, E.E.; Battey, J.; Ney, R.; Kirsch, I.R.; Leder, P.
Cell 29, 691-699, 1982
A/Title: Duplication and deletion in the human immunoglobulin epsilon genes.
A/Reference number: A90824; MUID:83001945; PMID:6286268
A/Accession: A90824
A/Molecule type: DNA
A/Residues: 1-358, 'L', 360-428 <MAX>
A/Cross-references: GB:J00222; NID:g184755
A/Note: This sequence difference may be due to polymorphism
R/Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.
In Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-3
A/Reference number: A94418
A/Accession: A94418
A/Molecule type: protein
A/Residues: 'GAWTL', 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 12
A/Experimental source: myeloma protein Nd

R/Kenten, J.H.; Molgaard, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.; G
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
A/Title: Cloning and sequence determination of the gene for the human immunoglobulin eps
A/Reference number: A93933; MUID:83065234; PMID:6815656
A/Accession: B93933
A/Molecule type: mRNA
A/Residues: 1-40;68-114;427-428 <KEN>
A/Cross-references: GB:L00022; NID:g185035
R/Ikeyama, S.
FEBS Lett. 224, 306-310, 1987
A/Title: Purification and characterization of a recombinant human IgE Fe-epsilon fragmen
A/Reference number: S02438; MUID:88083554; PMID:3121387
A/Accession: S02438
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 98-352 <IKE>
R/Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.
J. Biol. Chem. 269, 456-462, 1994
A/Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces
A/Reference number: A53116; MUID:94103254; PMID:8276835
A/Accession: A53116
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 320-428 <ZH2>
A/Experimental source: myeloma U266-derived cell line AF-10
A/Note: sequence extracted from NCBI backbone (NCBIN:141701, NCBIPI:141702)
R/Hellman, L.
Eur. J. Immunol. 23, 159-167, 1993
A/Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of
A/Reference number: A46536; MUID:93122085; PMID:8419166
A/Accession: A46536
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 382-426 <HEL>
A/Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167
A/Experimental source: B cell myeloma U-266
A/Note: sequence extracted from NCBI backbone (NCBIPI:125299)
A/Accession: A46536
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 382-391 <HE2>
A/Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169
A/Experimental source: B cell myeloma U-266
A/Note: sequence extracted from NCBI backbone (NCBIPI:125299)
A/Accession: A46536
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 401-428 <HE3>
A/Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A/Experimental source: B cell myeloma U-266
A/Note: sequence extracted from NCBI backbone (NCBIPI:123483)
C/Genetics:
A/Gene: GDB:IGHE
A/Cross-references: GDB:119335; OMIM:147180
A/Map position: 14q32.33-14q32.33
A/Introns: 1/1; 104/1; 211/1; 319/1
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C/Superfamily: immunoglobulin C region; immunoglobulin homology
F/22-87/Domain: immunoglobulin homology <IM1>
F/128-195/Domain: immunoglobulin homology <IM2>
F/232-301/Domain: immunoglobulin homology <IM3>
F/338-407/Domain: immunoglobulin homology <IM4>
F/14/Disulfide bonds: interchain (to light chain) #status predicted
F/15-105,29-85,135-193,239-299,345-405/Disulfide bonds: #status predicted
F/21,49,99,146,252,275/Binding site: carbohydrate (Asn) (covalent) #status experimental
F/121,209/Disulfide bonds: interchain (to heavy chain) #status predicted
Query Match 40.1%; Score 745; DB 1; Length 428;
Best Local Similarity 44.4%; Pred. No. 1.5e-47;
Matches 143; Conservative 62; Mismatches 109; Indels 8; Gaps 5;

```
QY 22 PPTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTVLWDGQAEENLFPYTTTRPKREG 81
Db 111 PPTVKILQSSCDGGGHPPTIQLLCLVSGYTPGTINITWLEDQVMD--VDLSTASTTQ 168
QY 82 GQTFSLQSEVNITQGWMSNTYTVCHVKHNGSIFEDSSRRCSDDPRGVITYLIPPSPLD 141
Db 169 GELASTQSELTLSQKHWSLDRITYTCQVYQGTTFEDSTKKCADSNPRGVSAVLSRSPFD 228
QY 142 LVYENGTPKLTCLVLDL-ESEENITVWVRERKKSIGSASQSTKHHATTSTSLPVD 200
Db 229 LPIKSPITITCLVLDLAPSGKTVNUTWSRASKPVNHSTRKEEKQKORGLTITVSTLPVGT 288
QY 201 KDWIEGEGYQCRVDHPHFKPIVRSITKLPGKRLAPEVYMLPPSPETGT--TRTVTCL 258
Db 289 RDWIEGETYQCRVTHPHLPALMRSTTKTSGPRAAPEVYAF-ATPEWPGSRDKRTLACLI 347
QY 259 RGFYSEISVQWLPNNEEDHTGHTTTPQKHGTDPSFFLYSRMLVKNKSWEKGNLVT 318
Db 348 QNFMPEDISVQWLHNEVQLPDARHSTTQPKTKGS--GFFVFSRLLEVTRAWEQKDEFIC 405
QY 319 RVVHEALPGSRTLEKSLHYSAG 340
Db 406 RAVHEAASPSQTVQRAVSVPNG 427
```

RESULT 6

I36948

Ig epsilon-chain - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C:Accession: I36948
R:Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orangutan
A:Reference number: I36948
A:Accession: I36948
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <RES>
A:Cross-references: GB:M15398; NID:g176797; PIDN:AAA35416.1; PID:g176798
C:Genetics:
A:introns: 103/1; 209/1; 317/1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 39.9%; Score 741; DB 2; Length 426;
Best Local Similarity 44.5%; Pred. No. 2.9e-47;
Matches 143; Conservative 60; Mismatches 110; Indels 8; Gaps 5;

```
QY 23 PTVKLFHSSCDPRGDAHSTIQLLCLVSGFSPAKVHVTVLWDGQAEENLFPYTTTRPKREG 82
Db 110 PTVKVLQSSCDGGGHPPTIQLLCLVSGYTPGTINITWLEDQVMD--VDLSTASATQEG 167
QY 83 QTFSLQSEVNITQGWMSNTYTVCHVKHNGSIFEDSSRRCSDDPRGVITYLIPPSPLD 142
Db 168 ELASTQSELTLSQKHWSLDRITYTCQVYQGTTFEDSTKKCADSNPRGVSAVLSRSPFD 227
QY 143 YENGTPKLTCLVLDL-ESEENITVWVRERKKSIGSASQSTKHHATTSTSLPVD 201
Db 228 FIRKSPITITCLVLDLAPSGKTVNUTWSRASKPVNHSTRKEEKQKORGLTITVSTLPVGT 287
QY 202 DWIEGEGYQCRVDHPHFKPIVRSITKLPGKRLAPEVYMLPPSPETGT--TRTVTCLIR 259
Db 288 DWIEGETYQCRVTHPHLPALMRSTTKTSGPRAAPEVYAF-ATPEGSGSRDKRTLACLI 346
QY 260 GYPSEISVQWLPNNEEDHTGHTTTPQKHGTDPSFFLYSRMLVKNKSWEKGNLVT 319
Db 347 NFMPEDISVQWLHNEVQLPDARHSTTQPKTKGS--GFFVFSRLLEVTRAWEQKDEFIC 404
QY 320 VVHEALPGSRTLEKSLHYSAG 340
Db 405 AVHEAASPSQTVQRAVSVPNG 425
```

RESULT 7

B46529

Ig Y heavy chain (7.8S) - duck
N:Alternate names: Ig gamma chain (7.8S)
C:Species: Anas platyrhynchos (domestic duck)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: B46529; S20759
R:Magor, K.E.; Warr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.
J. Immunol. 149, 2627-2633, 1992
A:Title: Structural relationship between the two IgY of the duck, Anas platyrhynchos: mo
A:Reference number: A46529; MUID:93017865; PMID:1401901
A:Accession: B46529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-572 <MAG>
A:Cross-references: EMBL:X65219; NID:g62442; PIDN:CAA46322.1; PID:g62443
A:Experimental source: spleen
A:Note: sequence extracted from NCBI backbone (NCBIP:116127)
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:37-120/Domain: immunoglobulin homology <IMM>

Query Match 31.6%; Score 587; DB 2; Length 572;

Best Local Similarity 37.3%; Pred. No. 9.8e-36;
Matches 121; Conservative 60; Mismatches 129; Indels 14; Gaps 8;

```
QY 22 PPTVKLFHSS-CDPRGDAHSTIQLLCLVSGFSPAKVHVTVLWDGQAEENLFPYTTTRPKRE 80
Db 249 PPEVQVHLVSCVSTLGD--DSVELLCVITGSPPPVEVWLVDGAPA-HLVATWTRPQRE 305
QY 81 -GGQTFSLQSEVNITQGWMSNTYTVCHVKH--NGSIFEDSSRRCSDDPRG---VITYL 134
Db 306 AGSKTYMATQTVNSREDWAKAGFTCRVKHPATGGTAQGHAFPCPSGAGSCSPQIFV 365
QY 135 IPPSPDLYENGTPKLTCLVLDLSEENITVWVRERKKSIGSASQSTKHHATTSTITS 194
Db 366 VPPSPGSLYIRQDAKVHCLVNLPSDASLSISWTRKSGALRPDPMLVTEHFNCTFTASS 425
QY 195 ILPVDADKWTGEGYQCRVDHPHFKPIVRSITKLPGKRLAPEVYMLPPSPETGTTR-T 253
Db 426 SLAISTQDLAGERFTCTVQHEDLFVPLGKSIAGHAKGVATPYITFPFHAEELSLAEVT 485
QY 254 VTCLIRGFYSEISVQWLPNNEEDHTGHTTTPQKHGTDPSFFLYSRMLVKNKSWEKG 313
Db 486 LTCLVRGFQPEHVEVQWLNRHNSVPAAEFVTTPLKEPNGDGTFFLYSKMTVPKASWQGG 545
QY 314 NLVTCRVVHEALP---GSRILEKS 334
Db 546 VSYACMVVHEGLPMRFTQRPLOKT 569
```

RESULT 8

S00390

Ig gamma chain (clone 36) - chicken (fragment)
N:Alternate names: Ig nu chain
C:Species: Gallus gallus (chicken)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C:Accession: S00390
R:Parvari, R.; Avivi, A.; Leshner, F.; Ziv, E.; Tel-Or, S.; Burstein, Y.; Schechter, I.
EMBO J. 7, 739-744, 1988
A:Title: Chicken immunoglobulin gamma-heavy chains: limited VH gene repertoire, combinat
A:Reference number: S00390; MUID:88283642; PMID:3135182
A:Accession: S00390
A:Molecule type: mRNA
A:Residues: 1-504 <PAR>
A:Cross-references: EMBL:X07174
A:Note: this sequence was determined from the differentiated gene
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match 30.1%; Score 560; DB 2; Length 504;

Best Local Similarity 35.7%; Pred. No. 8.2e-34;

A>Title: Comparison of complete nucleotide sequence of the human IgM heavy chain constant region
 A/Reference number: S37767; MUID:93109369; PMID:8417370
 A/Accession: S37768
 A/Molecule type: mRNA
 A/Residues: 1-453 <HAR>
 A/Cross-references: EMBL:X67301; NID:g38407; PIDN:CAA7714.1; PID:g38408
 A/Experimental source: cell line Ab 63
 C/Genetics:

A/Map position: 14q32
 C/Superfamily: immunoglobulin C region; immunoglobulin homology
 C/Keywords: glycoprotein; heterotetramer; immunoglobulin
 F:121-90/Domain: immunoglobulin homology <IMM1>
 F:127-199/Domain: immunoglobulin homology <IMM2>
 F:237-305/Domain: immunoglobulin homology <IMM3>
 F:344-415/Domain: immunoglobulin homology <IMM4>
 F:14/Disulfide bonds: interchain (to light chain) #status predicted
 F:28-88,134-197,244-303,351-413/Disulfide bonds: #status predicted
 F:46,209,272,279,440/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:214,452/Disulfide bonds: interchain (to heavy chain) #status predicted
 F:291/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 26.3%; Score 488.5; DB 2; Length 453;
 Best Local Similarity 31.2%; Pred. No. 1.4e-28;
 Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

```

Qy 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHVTLWLDGQAEAN 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 PLPVIAELPPKVSF---VPRDGFNGPRKS-KLICATGFSPRQIQVSLREGKQVGS 157

Qy 70 LFPYTT-----RPKREGQTSLOSEVNITQGMSSNTYCHVKHNGSIF-EDSSRRCS 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 --GVTTDQVQAEAKESGPTTKYSTLTIKESDMLSQSMFTCRVDHRLGLTFQONASSMCV 215

Qy 124 DDEPRGVITYLIPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKKSIGASQSRST 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 PDQTAIRVAIPSPFASIFLTKSTKLTCLVLDLTLYDSVTISWTRQNGEAVKTHTNISE 275

Qy 184 KHHHATTISITLIPVDAKDWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APEVYMLP 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 276 SHPNATFSAVGEASICEDDWNSEGFCTVTHDLPSPKQTSIRPKGVALHRDVLPL 335

Qy 243 PSPEETG--TTRVTCLIRGYPSISVQWLPNNEEDHTGHTTTRPKOHDGTDPSFPLY 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 PAREQLNLRRESATITCLVTGSPADVFQVMQGOPLSPKPYVTSAPMPPEQAPGRYFAH 395

Qy 301 SRMLVNKSIWEKGNLVTCTRVVHEALPG---SRTLEKS 334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 SILTVSEEWNTGETYTCVVAHEALPNRVTERTVDKS 432
  
```

RESULT 12

Ig heavy chain - human
 C/Species: Homo sapiens (man)
 C/Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 C/Accession: S15590
 R/Neale, G.A.M.; Kitchingman, G.R.
 Nucleic Acids Res. 19, 2427-2433, 1991
 A>Title: mRNA transcripts initiating within the human immunoglobulin mu heavy chain enhancer region
 A/Reference number: S15590; MUID:91252286; PMID:1904154
 A/Accession: S15590
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-474 <NEA>

A/Cross-references: EMBL:X58529
 A/Note: the authors translated the codon CAA for residue 265 as Glu
 C/Superfamily: immunoglobulin C region; immunoglobulin homology
 C/Keywords: heterotetramer; immunoglobulin
 F:344-415/Domain: immunoglobulin homology <IMM>

Query Match 26.3%; Score 488.5; DB 2; Length 474;
 Best Local Similarity 31.2%; Pred. No. 1.5e-28;
 Matches 105; Conservative 73; Mismatches 134; Indels 25; Gaps 10;

```

Qy 17 PVTII---PPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHVTLWLDGQAEAN 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 PLPVIAELPPKVSF---VPRDGFNGPRKS-KLICATGFSPRQIQVSLREGKQVGS 157

Qy 70 LFPYTT-----RPKREGQTSLOSEVNITQGMSSNTYCHVKHNGSIF-EDSSRRCS 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 --GVTTDQVQAEAKESGPTTKYSTLTIKESDMLSQSMFTCRVDHRLGLTFQONASSMCV 215

Qy 124 DDEPRGVITYLIPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKKSIGASQSRST 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 PDQTAIRVAIPSPFASIFLTKSTKLTCLVLDLTLYDSVTISWTRQNGEAVKTHTNISE 275

Qy 184 KHHHATTISITLIPVDAKDWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APEVYMLP 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 276 SHPNATFSAVGEASICEDDWNSEGFCTVTHDLPSPKQTSIRPKGVALHRDVLPL 335

Qy 243 PSPEETG--TTRVTCLIRGYPSISVQWLPNNEEDHTGHTTTRPKOHDGTDPSFPLY 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 PAREQLNLRRESATITCLVTGSPADVFQVMQGOPLSPKPYVTSAPMPPEQAPGRYFAH 395

Qy 301 SRMLVNKSIWEKGNLVTCTRVVHEALPG---SRTLEKS 334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 SILTVSEEWNTGETYTCVVAHEALPNRVTERTVDKS 432
  
```

RESULT 13

MRB

Ig mu chain C region, secreted form - rabbit
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 23-Aug-1997
 C/Accession: A02164
 J/Bernstein, K.E.; Alexander, C.B.; Reddy, E.P.; Mage, R.G.
 J. Immunol. 132, 490-495, 1984
 A>Title: Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of V-H2a 8
 A/Reference number: A02164; MUID:84089930; PMID:6418803
 A/Contents: a2 allotype
 A/Accession: A02164
 A/Molecule type: mRNA
 A/Residues: 1-458 <BER>

C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la C/Superfamily: immunoglobulin C region; immunoglobulin homology
 C/Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin
 F:21-92/Domain: immunoglobulin homology <IMM1>
 F:130-202/Domain: immunoglobulin homology <IMM2>
 F:242-310/Domain: immunoglobulin homology <IMM3>
 F:349-420/Domain: immunoglobulin homology <IMM4>
 F:14/Disulfide bonds: interchain (to light chain) #status predicted
 F:28-90,137-200,249-308,356-418/Disulfide bonds: #status predicted
 F:46,114,212,261,277,284,445/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:219,457/Disulfide bonds: interchain (to heavy chain) #status predicted
 F:296/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 26.3%; Score 488; DB 1; Length 458;
 Best Local Similarity 32.7%; Pred. No. 1.5e-28;
 Matches 114; Conservative 62; Mismatches 149; Indels 24; Gaps 9;

```

Qy 5 HHHHTLSLSPSGPV-TIIPVTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHV 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 HSNRNRLRVSPFVDPSELPPNVSVF---IPPRDFSGSGTRKSLICQATGFSFKQISVS 150

Qy 60 WLVDGQAEAN--LFPVTTTRPKREGGOTFSLQSEVNITQGMSSNTYCHVKHNGSIFE 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 WLDRGQKVESGLVTKPVEAETKAGPATFSSMLTITSDWLSQSILYTCVDRHGIFFD 210

Qy 117 DS---SRRCSDEPRGVITYLIPSPDLVYENGTPKLTCLVLDLESEENITVTVRERKK 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 KNVSMSECSCTTSPGIVQVFPAPSPADTFLSKSARLCLVLDLTLYGSLNLSWASHNGK 270

Qy 174 SIGSASORSTKHHTTISITLIPVDAKDWIEGEGYQCRVDHPHPKPIVRSITKLPKR 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 ALDTHMNITESHNPATFSAMGEASVCAEDWESGEQFTCTVTTHADLPFLKHTISK--SRE 328
  
```


Search completed: July 9, 2003, 13:39:44
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 13:39:16 ; Search time 15 Seconds
(without alignments)
668.881 Million cell updates/sec

Title: US-09-401-636-4

Perfect score: 1858

Sequence: 1 EFHHHHHTLSLPESGPVTI.....HEALPGSRTEKSLHYSAGN 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	950.5	51.2	426	US-08-336-583-2	Sequence 2, Appli
2	950.5	51.2	426	PCT-US95-13795-2	Sequence 2, Appli
3	862.5	46.4	561	US-09-192-545-2	Sequence 2, Appli
4	507	27.3	331	US-08-646-981-17	Sequence 17, Appl
5	493.5	26.6	334	US-08-646-981-16	Sequence 16, Appl
6	488	26.3	504	US-07-932-915-2	Sequence 2, Appli
7	488	26.3	504	PCT-US91-05826-2	Sequence 2, Appli
8	475.5	25.6	333	US-08-436-453-6	Sequence 6, Appli
9	475.5	25.6	333	US-08-024-253-6	Sequence 6, Appli
10	475	25.6	450	US-08-788-800-12	Sequence 12, Appl
11	475	25.6	469	US-07-934-373C-23	Sequence 23, Appl
12	475	25.6	469	US-08-437-642B-23	Sequence 23, Appl
13	475	25.6	469	US-08-146-206C-23	Sequence 23, Appl
14	469.5	25.3	530	US-08-477-450B-4	Sequence 4, Appli
15	469.5	25.3	530	US-08-379-516-4	Sequence 4, Appli
16	469.5	25.3	530	US-09-329-916-4	Sequence 4, Appli
17	469.5	25.3	530	US-08-485-372A-4	Sequence 4, Appli
18	469.5	25.3	530	US-09-409-006A-4	Sequence 4, Appli
19	469.5	25.3	530	US-08-484-681-4	Sequence 4, Appli
20	469.5	25.3	530	PCT-US93-07422-4	Sequence 4, Appli
21	457	24.6	326	US-08-656-586-9	Sequence 9, Appli
22	454	24.4	552	PCT-US93-07832-23	Sequence 23, Appl
23	451.5	24.3	327	US-08-761-277A-47	Sequence 47, Appl
24	451.5	24.3	443	PCT-US96-13152-4	Sequence 4, Appli
25	451.5	24.3	467	US-08-704-744-81	Sequence 81, Appl
26	451.5	24.3	467	US-08-523-894-8	Sequence 8, Appli
27	451.5	24.3	467	US-08-523-894-10	Sequence 10, Appl

28	451.5	24.3	467	4	US-08-523-894-12	Sequence 12, Appl
29	447	24.1	451	4	US-09-247-352-3	Sequence 3, Appli
30	447	24.1	451	4	US-09-466-635-3	Sequence 3, Appli
31	447	24.1	472	4	US-09-301-593-30	Sequence 30, Appl
32	447	24.1	472	4	US-09-301-593-43	Sequence 43, Appl
33	446.5	24.0	451	2	US-08-887-352B-14	Sequence 14, Appl
34	446.5	24.0	451	2	US-08-887-352B-16	Sequence 16, Appl
35	446.5	24.0	451	2	US-08-887-352B-18	Sequence 18, Appl
36	446.5	24.0	451	3	US-08-466-151-65	Sequence 65, Appl
37	446.5	24.0	451	4	US-09-109-207C-14	Sequence 14, Appl
38	446.5	24.0	451	4	US-09-109-207C-16	Sequence 16, Appl
39	446.5	24.0	451	4	US-09-109-207C-18	Sequence 18, Appl
40	446.5	24.0	451	4	US-09-282-505-2	Sequence 2, Appli
41	446.5	24.0	451	4	US-09-054-255-2	Sequence 2, Appli
42	446.5	24.0	451	4	US-09-296-005-14	Sequence 14, Appl
43	446.5	24.0	451	4	US-09-296-005-16	Sequence 16, Appl
44	446.5	24.0	451	4	US-09-296-005-18	Sequence 18, Appl
45	445.5	24.0	453	3	US-08-466-151-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-336-583-2
; Sequence 2, Application US/08336583
; Patent No. 5629415
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,583
; FILING DATE: 09-NOV-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-583-2

Query Match 51.2%; Score 950.5; DB 1; Length 426;
Best Local Similarity 53.8%; Pred. No. 2.2e-86;
Matches 175; Conservative 52; Mismatches 95; Indels 3; Gaps 2;

Oy	18	VTIPTVKLFHSSCDPRGDAHSTIQLCLVSGSPAKVHVTVLVDGQAEALPPYTRP	77
Db	102	LNFIPTVKLFHSSCNVPGDTHITQLCLISGVPGDMEVILVDGQAKINIFPYAPG	161
Oy	78	KREGQTFLSQSEVNITQGMSSNTYTCHVKHNGSIFEDSSRRCSDDPRGVITYLIPP	137

Db 162 TKEGNTV-STHSELNITGEVWSQKTYTCQVYQGTFTKDEARKCSBSPRGVTSYLSPP 220
Qy 138 SPLDLYENGTPKLTCLVLDLSEENITVWVRERKKSIGSASQSRSTKHHHATTSTISILP 197
Db 221 SPLDLYVHKAPKITCLVVDLATMEGNLTWYRESKEPVNPGPLNKDKHFNIGTITVTSTLP 280
Qy 198 VDADWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYMLPPSPETGTTR--TVT 255
Db 281 VNTNDWIEGETYCKRTVPHLPKDIVRSIAKAPGRAPPDVYLFPPBEEQGTDRVTLT 340
Qy 198 VDADWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYMLPPSPETGTTR--TVT 255
Db 281 VNTNDWIEGETYCKRTVPHLPKDIVRSIAKAPGRAPPDVYLFPPBEEQGTDRVTLT 340
Qy 256 CLIRGFYPSSEISVQWLNNEDHTGHTTTRPQDGHGTDPSFFLYSRMLVNKSWEKGNL 315
Db 341 CLIQNFPPADISVQWLNRDPSIQDQTTTGPHKVGSRPAFFIFSRLEVRSDVWEQKNK 400
Qy 316 VTCRVVHEALPGSRILEKSLHYSAG 340
Db 401 FTCQVVEALSGSRILQKWSKTPG 425

RESULT 2

PCT-US95-13795-2
; Sequence 2, Application PC/TUS9513795
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE, P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13795
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19211Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-13795-2

Query Match 51.2%; Score 950.5; DB 5; Length 426;
Best Local Similarity 53.8%; Pred. No. 2.2e-86;
Matches 175; Conservative 52; Mismatches 95; Indels 3; Gaps 2;
Qy 18 VTIIPPTVKLFHSSCDPRGDHSTTQLCLVSGFSPAKVHVTLVVDGQAEENLFPYTRP 77
Db 102 LNFIPPTVKLFHSSCNVGDTHTTQLCLISGYVPGDMEVTLVVDGQATNIPFYTAGP 161
Qy 78 KREGQGTSLQSEVNITQGMSSNTYCHVKHNGSIFEDSSRRCSDDPRGVITYLIPP 137
Db 162 TKEGNTV-STHSELNITGEVWSQKTYTCQVYQGTFTKDEARKCSBSPRGVTSYLSPP 220
Qy 138 SPLDLYENGTPKLTCLVLDLSEENITVWVRERKKSIGSASQSRSTKHHHATTSTISILP 197

Db 221 SPLDLYVHKAPKITCLVVDLATMEGNLTWYRESKEPVNPGPLNKDKHFNIGTITVTSTLP 280
Qy 198 VDADWIEGEGYQCRVDHPHPKPIVRSITKLPGRKLAPEVYMLPPSPETGTTR--TVT 255
Db 281 VNTNDWIEGETYCKRTVPHLPKDIVRSIAKAPGRAPPDVYLFPPBEEQGTDRVTLT 340
Qy 256 CLIRGFYPSSEISVQWLNNEDHTGHTTTRPQDGHGTDPSFFLYSRMLVNKSWEKGNL 315
Db 341 CLIQNFPPADISVQWLNRDPSIQDQTTTGPHKVGSRPAFFIFSRLEVRSDVWEQKNK 400
Qy 316 VTCRVVHEALPGSRILEKSLHYSAG 340
Db 401 FTCQVVEALSGSRILQKWSKTPG 425
RESULT 3
US-09-192-545-2
; Sequence 2, Application US/09192545
; Patent No. 6118044
; GENERAL INFORMATION:
; APPLICANT: Karasuyama, Hajime
; APPLICANT: Yonekawa, Hiromichi
; APPLICANT: Taya, Choji
; APPLICANT: Matsuo, Kunie
; TITLE OF INVENTION: Transgenic Animal Allergy Models and Methods for Their Use
; CURRENT APPLICATION NUMBER: US/09/192,545
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: JP HEI 9-313989
; EARLIER FILING DATE: 1997-11-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: Designed heavy
US-09-192-545-2

Query Match 46.4%; Score 862.5; DB 3; Length 561;
Best Local Similarity 51.8%; Pred. No. 1.9e-77;
Matches 183; Conservative 46; Mismatches 103; Indels 21; Gaps 8;
Qy 2 FHHHHHTLSLPESGPVTIIPPTVKLFHS-----SCDPRGDAHSTIQLCLVSGP 51
Db 211 FTCHVTHPPSPFNSRSTILVRPVT---HSLSPPWSYSIHRCDPNA-FHSTIQLYCFYGH 265
Qy 52 SPKAVHVTVLVDGQAEENLFPYTRPKREGGQTFSLQSEVNITQGMSSNTYCHVKHN 111
Db 266 ILNDVSYSWLMDOREITDTLAQTVLKEE-GKLASTCSKLNITBQOMSESTFCRVTSQ 324
Qy 112 GSIFEDSSRRCSDDPRGVITYLIPPSPLDLYENGTPKLTCLVLDLSEENITVWVRER 171
Db 325 GVDYLAHTRRCPOHEPRGATYLIIPPSPLDLYQNGAPKLTCLVVDLSEKKNVNTWNOEK 384
Qy 172 KKSIGSASQSRSTKH-HATTSTISILPVDADWIEGEGYQCRVDHPHPKPIVRSITKLP 230
Db 385 KTSV-SASQWYTRHNNATTSITSLPVAKDWIEGYQCVDRPDPFPKPIVRSIT-LP 442
Qy 231 --GKRLAPEVYMLPPSPETGTTRTCLIRGFYPSSEISVQWLNNEDHTGHTTTRPQ 288
Db 443 QVSQSAPEVYVFPPEESEDKETLTCLIQNFPPEDISVQWLGDGKLIINSQHSTTTP 502
Qy 289 KDHGTDPSFFLYSRMLVNKSWEKGNLVTQVVEALPGSRILEKSLHYSAGN 341
Db 503 KSNQSGQGTFFSRLEVAKTWTQKQTCQVTHEALQKPRKLEKTIISTISGN 555
RESULT 4
US-08-646-981-17
; Sequence 17, Application US/08646981
; Patent No. 5852183


```

GENERAL INFORMATION:
APPLICANT: MAEDA, HIROAKI
APPLICANT: EDA, YASUYUKI
APPLICANT: KIMACHI, KAZUHIKO
APPLICANT: ONO, YOICHI
APPLICANT: TOKIYOSHI, SACHIO
TITLE OF INVENTION: DOG-MOUSE HETEROHYBRIDOMA AND GENE
TITLE OF INVENTION: FRAGMENT CODING FOR CONSTANT REGION OF CANINE
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,981
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 1488-106
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 331 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-646-981-17

Query Match          27.3%   Score 507;   DB 2;   Length 331;
Best Local Similarity 35.2%;   Pred. No. 2.3e-42;
Matches 118;   Conservative 69;   Mismatches 122;   Indels 26;   Gaps 11;

QY      23      PTYKLFHSCDPRGDAHSTIQLCLVSGSPAKVHVHTLWLDGQEAENLFPYITRPREGG 82
Db      5      PSVFPLAPSCG--STGSGTTLACLVGSGYFPEPTVSW-NSGLTSGVHTFPVLKSSG- 60
QY      83      QTESLQSEVNITOGWMSNTYICHVKNGS-----TFEDSRRCSDDEP-----RG 129
Db      61      -LVSLSGSMVTVPSSR-LPSEFTFCNVVHPATNTKVDKPGVPKSTCKISCPVPESLGG 118
QY      130     VITYLPPSPDLIYE-NGTEPKLTCLVLDLSEEE-NITVTWVRERKKSTGASQSRSTKHHH 187
Db      119     PSVEIPIPPPKDILIRTRTEVTCVLDLGREDEPVOISNFWDCGEVHTAKTQPREQQFN 178
QY      188     ATTSITSILPVDKADWJTEGGYQCRVDHPHPKPIVRSITKLPKRIAPAEVYMLPPSPPE 247
Db      179     STYRVVSVLPDIEHQDWTLTGKEFKCRVNHGLPSPDIERTISKARQAQHPGVYVLPSPKE 238
QY      248     --TGTTTRTVCLIRGFYPSISVOWLNNEEDHTGHHTTTRPOKDHGTDPSEFFLYSRMLV 305
Db      239     LSSSDTVTLTCLLNKDFPFPIIDVWQNGQPEPESEKTHHTTAPQLDE--DGSYFLYKLSV 296
QY      306     NKSIEWEKNGLVTRCVVHEALPGSRTEKLSHYSAG 340
Db      297     DKSRWEGDDEPTCAVMHEALONHYT-DLSLSHSPG 330

```

APPLICANT: MAEDA, HIROAKI
 APPLICANT: ED, YASUYUKI
 APPLICANT: KIMACHI, KAZUHIKO
 APPLICANT: ONO, YOICHI
 APPLICANT: TOKIYOSHI, SACHIO
 TITLE OF INVENTION: DOG-MOUSE HETEROHYBRIDOMA AND GENE
 TITLE OF INVENTION: FRAGMENT CODING FOR CONSTANT REGION OF CANINE
 TITLE OF INVENTION: IMMUNOGLOBULINS
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
 STREET: PO BOX 747
 CITY: FALLS CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,981
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: WEINER, MARC S
 REGISTRATION NUMBER: 32,181
 REFERENCE/DOCKET NUMBER: 1488-106
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 334 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-646-981-16

RESULT 6
US-07-932-915-2
; Sequence 2, Application US/07932915
; Patent No. 5672486
; GENERAL INFORMATION:
; APPLICANT: Soullilou, Jean-Paul

TITLE OF INVENTION: Protein Polyligands Joined To A Stable Protein
TITLE OF INVENTION: Core
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
STREET: 5 Palo Alto Square, Suite 400
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/932,915
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/646,875
FILING DATE: 28-JAN-1991
APPLICATION NUMBER: US 07/575,394
FILING DATE: 23-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Rowland Ph.D., Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: ATLA-001/0115
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-932-915-2

Query Match 26.3%; Score 488; DB 1; Length 504;
Best Local Similarity 31.2%; Pred. No. 3.5e-40;
Matches 103; Conservative 72; Mismatches 133; Indels 22; Gaps 9;
Qy 21 IPPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHVTVLVDGQAEENLFPVTT- 75
Db 160 LPPKVSVP---VPPRDGFGNPKRS-KLICATGFSRQIQVSMLEKQVGS--GVTTD 213
Qy 76 ---RPRKGGQTFSLQSEVNITQGMWSSNTYTVCHVKHNGSIF-EDSSRRCSDDPRGV 130
Db 214 QVQAEAKESGPTYKVTSTLTIKESDMLSQSMFTCRVDHRLGLTFOQNASSMCPVDDTAI 273
Qy 131 ITYLIPSPDLVYENGTPKLTCLVLDLSEENITVTVWRERKKSIGSASQSTKHHATT 190
Db 274 RVFAIPSPFASIFLTKSLKLTCLVTLDTLTYDSVTISWTRQNGEAVKTHNISESHPNATF 333
Qy 191 SITSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APVYMLPSPSPETG 249
Db 334 SAVGEASICEDDWSNGERFTCTVHTDLPSPKQIISRPKGVALHRPDVYLLPAREQLN 393
Qy 250 ---TTRTVTCLIRGFYPSISVQWLPNNEDHTGHHHTTRPKQDHGTDPSEFLYSRMLVNK 307
Db 394 LRESATITCLVTGSPADVQVQWQMGQPLSPEKYVTSAPMPPEQAPGRYFAHSILTYSVE 453
Qy 308 SIWEKGNLVTCTRVVHEALPG---SRTLEKS 334
Db 454 BEWNTGETYTCVVAHEALPNRVTERTVDKS 483

RESULT 7
PCT-US91-05826-2
Sequence 2, Application PC/TUS9105826
GENERAL INFORMATION:
APPLICANT: Soullilou, Jean-Paul

TITLE OF INVENTION: Protein Polyligands Joined To A Stable Protein
TITLE OF INVENTION: Core
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bertram I., Rowland, Ph.D.
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05826
FILING DATE: 19910822
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/575,394
FILING DATE: 29-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Rowland Ph.D., Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: PP55352-1/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-494-8771
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-05826-2

Query Match 26.3%; Score 488; DB 5; Length 504;
Best Local Similarity 31.2%; Pred. No. 3.5e-40;
Matches 103; Conservative 72; Mismatches 133; Indels 22; Gaps 9;
Qy 21 IPPTVKLFHSSCDPR---GDAHSTIQLCLVSGFSPAKVHVTVLVDGQAEENLFPVTT- 75
Db 160 LPPKVSVP---VPPRDGFGNPKRS-KLICATGFSRQIQVSMLEKQVGS--GVTTD 213
Qy 76 ---RPRKGGQTFSLQSEVNITQGMWSSNTYTVCHVKHNGSIF-EDSSRRCSDDPRGV 130
Db 214 QVQAEAKESGPTYKVTSTLTIKESDMLSQSMFTCRVDHRLGLTFOQNASSMCPVDDTAI 273
Qy 131 ITYLIPSPDLVYENGTPKLTCLVLDLSEENITVTVWRERKKSIGSASQSTKHHATT 190
Db 274 RVFAIPSPFASIFLTKSLKLTCLVTLDTLTYDSVTISWTRQNGEAVKTHNISESHPNATF 333
Qy 191 SITSLPVDADKWIEGEGYQCRVDHPHPKPIVRSITKLPKRL-APVYMLPSPSPETG 249
Db 334 SAVGEASICEDDWSNGERFTCTVHTDLPSPKQIISRPKGVALHRPDVYLLPAREQLN 393
Qy 250 ---TTRTVTCLIRGFYPSISVQWLPNNEDHTGHHHTTRPKQDHGTDPSEFLYSRMLVNK 307
Db 394 LRESATITCLVTGSPADVQVQWQMGQPLSPEKYVTSAPMPPEQAPGRYFAHSILTYSVE 453
Qy 308 SIWEKGNLVTCTRVVHEALPG---SRTLEKS 334
Db 454 BEWNTGETYTCVVAHEALPNRVTERTVDKS 483

RESULT 8
US-08-436-463-6
Sequence 6, Application US/08436463
Patent No. 5760185
GENERAL INFORMATION:
APPLICANT: KIMACHI, Kazuhiko
APPLICANT: MAEDA, Hiroaki

APPLICANT: NISHIYAMA, Kiyoto
APPLICANT: TOKIYOSHI, Sachio
TITLE OF INVENTION: ANTI-FELINE HERPES VIRUS-1 RECOMBINANT
TITLE OF INVENTION: ANTIBODY AND GENE FRAGMENT CODING FOR SAID ANTIBODY
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436.463
FILING DATE: 26-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 341255/1992
FILING DATE: 28-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: KIMACHI=1
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-436-463-6

Query Match 25.6%; Score 475.5; DB 1; Length 333;
Best Local Similarity 32.7%; Pred No. 3.3e-39;
Matches 112; Conservative 72; Mismatches 121; Indels 37; Gaps 12;
Qy 23 PTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVTVW----LVDGQAEANLFPYTRPK 78
Db 4 PSVPPLAPSCGTTSGA--TVALACLVLGYFPPEVTVSWNSGALTSG---VHTFPAVL--- 55
Qy 79 REGGQTFSLQSEVNITOGQWSSNTYTCVKHNGSIFE-DSSRRCSDDP----- 127
Db 56 -QASGLYSLSMVTVPSRWL-SDTFTCNVAHPSPNTKVDKTVRKTDHPGPKPCDCPKC 113
Qy 128 -----RGVITYLIPPSPLD-LYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGSAS 179
Db 114 PPPEMLGSPSIFIPPKPKOTLSISRTPEVTVCLVVDLGPDDSDVOITWFDVNTQVYTAKT 173
Qy 180 QRSTKHHATTSTISILPVDADKMWIEGEGYQCRVDHPFPKPIVRSITKLPKGLAPEVY 239
Db 174 SPREEQFNSTVRVSVLPILHQDLKGFCKKVNKSLPSPIERTISKAKGPHEPQVY 233
Qy 240 MLPPSPETGTTR-TVTCLIRGFYPSISVQWLPNNEEDHTGHTTTPQKHGTDPSFF 298
Db 234 VLPPAQEELSRNKVSVTCLIKSPHPPDIAVEWEITGQPEPNNTVTPPQLD--SDGYTF 291
Qy 299 LYSRMLVNKSIWEKGNLVTCTRVVHEALPGSRRTLEKSLHYSAG 340
Db 292 VYSKLSVDRSHWRGNTYTCVSHEALHSHHT-QKSLTQSPG 332

RESULT 9
US-08-024-253-6
Sequence 6, Application US/08024253
Patent No. 5785968
GENERAL INFORMATION:

APPLICANT: KIMACHI, Kazuhiko
APPLICANT: MAEDA, Hiroaki
APPLICANT: NISHIYAMA, Kiyoto
APPLICANT: TOKIYOSHI, Sachio
APPLICANT: TOHYA, Yukinobu
APPLICANT: MIKAMI, Takeshi
TITLE OF INVENTION: ANTI-FELINE CALCIVIRUS RECOMBINANT
TITLE OF INVENTION: ANTIBODY AND GENE FRAGMENT ENCODING THE SAME
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEGNER, CANTOR, MUELLER & PLAYER
STREET: 1233 20th Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036-8218
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/024.253
FILING DATE: 19930301
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 79189/1992
FILING DATE: 28-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: CANTOR, Herbert I.
REGISTRATION NUMBER: 24,392
REFERENCE/DOCKET NUMBER: P-500-23744
TELEPHONE: (202) 887-0400
TELEFAX: (202) 835-0605
TELEX: 440706 WEGBR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-024-253-6

Query Match 25.6%; Score 475.5; DB 1; Length 333;
Best Local Similarity 32.7%; Pred No. 3.3e-39;
Matches 112; Conservative 72; Mismatches 121; Indels 37; Gaps 12;
Qy 23 PTVKLFHSSCDPRGDAHSTIQLCLVSGFSPAKVHVTVW----LVDGQAEANLFPYTRPK 78
Db 4 PSVPPLAPSCGTTSGA--TVALACLVLGYFPPEVTVSWNSGALTSG---VHTFPAVL--- 55
Qy 79 REGGQTFSLQSEVNITOGQWSSNTYTCVKHNGSIFE-DSSRRCSDDP----- 127
Db 56 -QASGLYSLSMVTVPSRWL-SDTFTCNVAHPSPNTKVDKTVRKTDHPGPKPCDCPKC 113
Qy 128 -----RGVITYLIPPSPLD-LYENGTPKLTCLVLDL-ESEENITVTVWRERKKSIGSAS 179
Db 114 PPPEMLGSPSIFIPPKPKOTLSISRTPEVTVCLVVDLGPDDSDVOITWFDVNTQVYTAKT 173
Qy 180 QRSTKHHATTSTISILPVDADKMWIEGEGYQCRVDHPFPKPIVRSITKLPKGLAPEVY 239
Db 174 SPREEQFNSTVRVSVLPILHQDLKGFCKKVNKSLPSPIERTISKAKGPHEPQVY 233
Qy 240 MLPPSPETGTTR-TVTCLIRGFYPSISVQWLPNNEEDHTGHTTTPQKHGTDPSFF 298
Db 234 VLPPAQEELSRNKVSVTCLIKSPHPPDIAVEWEITGQPEPNNTVTPPQLD--SDGYTF 291
Qy 299 LYSRMLVNKSIWEKGNLVTCTRVVHEALPGSRRTLEKSLHYSAG 340
Db 292 VYSKLSVDRSHWRGNTYTCVSHEALHSHHT-QKSLTQSPG 332

RESULT 10
US-08-788-800-12
; Sequence 12, Application US/08788800
; Patent No. 5914112
; GENERAL INFORMATION:
; APPLICANT: Bednar, Martin M.
; APPLICANT: Thomas, G. Roger
; APPLICANT: Gross, Cordell E.
; TITLE OF INVENTION: ANTI-CD18 ANTIBODIES IN STROKE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,800
; FILING DATE: 22-Jan-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P09871
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-788-800-12

Query Match 25.6%; Score 475; DB 2; Length 450;
Best Local Similarity 34.7%; Pred. No. 5.8e-39;
Matches 112; Conservative 66; Mismatches 115; Indels 30; Gaps 12;
Qy 35 RGDHSTIQLLCVSGFSPAKVHTW----LVGQEAENLPYTRPKREGGQTFSLQSE 90
Db 140 RSTSESTAALGCLVKDYFPEPVTVMNSGALTSG---VHTFPAVL-----QSSGLYSLSV 192
Qy 91 VNITQGMSSNTYTCVHKH---NGSIFEDSSRRCSDDP-----RGVITYLIPSP 140
Db 193 VVTSSNF-GTQYTCNVDRKPSNTKDKTVKCCVCEPCPPAPVAGPSVFLPPKPK 251
Qy 141 D-LYENGTPKLTCLVLDLESE-NITVTWVRERKKSIGSASQSTKHHHTTSITSLPV 198
Db 252 DTLMSIRTEPVCVVVDVSHEDPEVQFNWYVDGMEVHNAKTKPREEQNSFRVVSULTV 311
Qy 199 DAKWIEGEGYQCRVDHPFKPIVRSITKLPKGLAPEVVMPLPSPPEE-TGTRTIVTCL 257
Db 312 VHQDLWNGEKYCKVSNKGLPAPIEKTISKTKGQPREQVYTLPPSREEMTKNQVSLTCL 371
Qy 258 IRGVPSEISVOWLPNNEEDHTGHTTTRPKDGHGTPSPFLYSRLVKNKSIWEKGNLVT 317
Db 372 VKGFPSPDIJAVESNGQPEN--NYKTPPMLD---SDGSFFLYSLKLTVDKSRWQGNVFS 427
Qy 318 CRVVEALPGSRITLESLSHSAG 340
Db 428 CSVMHEALHNHYT-QKSLSLSPG 449

RESULT 11
US-07-934-373C-23
; Sequence 23, Application US/07934373C

Patent No. 5821337
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,373C
; FILING DATE: 21-Aug-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-23

Query Match 25.6%; Score 475; DB 2; Length 469;
Best Local Similarity 34.7%; Pred. No. 6.2e-39;
Matches 112; Conservative 66; Mismatches 115; Indels 30; Gaps 12;
Qy 35 RGDHSTIQLLCVSGFSPAKVHTW----LVGQEAENLPYTRPKREGGQTFSLQSE 90
Db 159 RSTSESTAALGCLVKDYFPEPVTVMNSGALTSG---VHTFPAVL-----QSSGLYSLSV 211
Qy 91 VNITQGMSSNTYTCVHKH---NGSIFEDSSRRCSDDP-----RGVITYLIPSP 140
Db 212 VVTSSNF-GTQYTCNVDRKPSNTKDKTVKCCVCEPCPPAPVAGPSVFLPPKPK 270
Qy 141 D-LYENGTPKLTCLVLDLESE-NITVTWVRERKKSIGSASQSTKHHHTTSITSLPV 198
Db 271 DTLMSIRTEPVCVVVDVSHEDPEVQFNWYVDGMEVHNAKTKPREEQNSFRVVSULTV 330
Qy 199 DAKWIEGEGYQCRVDHPFKPIVRSITKLPKGLAPEVVMPLPSPPEE-TGTRTIVTCL 257
Db 331 VHQDLWNGEKYCKVSNKGLPAPIEKTISKTKGQPREQVYTLPPSREEMTKNQVSLTCL 390
Qy 258 IRGVPSEISVOWLPNNEEDHTGHTTTRPKDGHGTPSPFLYSRLVKNKSIWEKGNLVT 317
Db 391 VKGFPSPDIJAVESNGQPEN--NYKTPPMLD---SDGSFFLYSLKLTVDKSRWQGNVFS 446
Qy 318 CRVVEALPGSRITLESLSHSAG 340
Db 447 CSVMHEALHNHYT-QKSLSLSPG 468

RESULT 12
US-08-437-642B-23

Sequence 23, Application US/08437642B
Patent No. 6054297
GENERAL INFORMATION:
APPLICANT: Paul J. Carter
APPLICANT: Leonard G. Presta
TITLE OF INVENTION: Immunoglobulin Variants
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/437,642B
FILING DATE: 09-May-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/934373
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/146206
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05126
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709P2C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-437-642B-23

Query Match 25.6%; Score 475; DB 3; Length 469;
Best Local Similarity 34.7%; Pred. No. 6.2e-39;
Matches 112; Conservative 66; Mismatches 115; Indels 30; Gaps 12;
Qy 35 RGDASHIQLLCVSGSPAKVHTW-----LVDQEAENLPYTRPKREGGQTFSLQSE 90
Db 159 RSTSESTAALGCLVKDYFPEPTVSNMGSALTSG---VHTPPAVL-----QSSGLYSLSV 211
Qy 91 VNITQGMWSNTVTCHVKH---NGSIFEDSSRRCSDDP-----RGVITYLIPPSPL 140
Db 212 VVTSSNP-GTQYTCNVNDHKPSNTKVDKVERKCCVCEPCPCAPPVAGPSVFLPPPKP 270
Qy 141 D-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRTSKHHATTSTISILPV 198
Db 271 DTLMSRTPETCVVDVSHEDPEVQFNWYVDGMEVHNNAKTKPREQFNSTFRVSVLTV 330
Qy 199 DAKWIEGEGYQCRVDHPHFKPIVRSITKUPGKRLAPEVYMLPPSPPEE-TGTRTWTCL 257
Db 331 VHODWLNKGYKCKVSNKGLPAPIETKISTKGOPREPQVYTLPPSREEMTKNOVSLTCL 390
Qy 258 IRGFYPSISVQWLPNNEEDHTGHTTTRPKQDGTDPSPFLYSRMLVNKSIWEKGNLVT 317
Db 391 VKGFYPSDIAVESWNGQPN--NYKTTTPMLD--SDGSFFLYSKLTVDKSRWQGNVFS 446

Qy 318 CRVVEALPGSRTLEKSLHYSAG 340
Db 447 CSVWHEALHNYT-QKSLSLSPG 468
RESULT 13
US-08-146-206C-23
Sequence 23, Application US/08146206C
Patent No. 6407213
GENERAL INFORMATION:
APPLICANT: Carter, Paul J.
APPLICANT: Presta, Leonard G.
TITLE OF INVENTION: Method for Making Humanized Antibodies
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,206C
FILING DATE: 17-No. 6407213-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/715272
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0709P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-146-206C-23

Query Match 25.6%; Score 475; DB 4; Length 469;
Best Local Similarity 34.7%; Pred. No. 6.2e-39;
Matches 112; Conservative 66; Mismatches 115; Indels 30; Gaps 12;
Qy 35 RGDASHIQLLCVSGSPAKVHTW-----LVDQEAENLPYTRPKREGGQTFSLQSE 90
Db 159 RSTSESTAALGCLVKDYFPEPTVSNMGSALTSG---VHTPPAVL-----QSSGLYSLSV 211
Qy 91 VNITQGMWSNTVTCHVKH---NGSIFEDSSRRCSDDP-----RGVITYLIPPSPL 140
Db 212 VVTSSNP-GTQYTCNVNDHKPSNTKVDKVERKCCVCEPCPCAPPVAGPSVFLPPPKP 270
Qy 141 D-LYENGTPKLTCLVLDLESEE-NITVTWVRERKKSIGSASQRTSKHHATTSTISILPV 198
Db 271 DTLMSRTPETCVVDVSHEDPEVQFNWYVDGMEVHNNAKTKPREQFNSTFRVSVLTV 330
Qy 199 DAKWIEGEGYQCRVDHPHFKPIVRSITKUPGKRLAPEVYMLPPSPPEE-TGTRTWTCL 257
Db 331 VHODWLNKGYKCKVSNKGLPAPIETKISTKGOPREPQVYTLPPSREEMTKNOVSLTCL 390
Qy 258 IRGFYPSISVQWLPNNEEDHTGHTTTRPKQDGTDPSPFLYSRMLVNKSIWEKGNLVT 317
Db 391 VKGFYPSDIAVESWNGQPN--NYKTTTPMLD--SDGSFFLYSKLTVDKSRWQGNVFS 446
Qy 318 CRVVEALPGSRTLEKSLHYSAG 340

Db 447 CSVMEALHNNHT-QKSLSPG 468

RESULT 14

US-08-477-460B-4

; Sequence 4, Application US/08477460B

; Patent No. 6034223

; GENERAL INFORMATION:

; APPLICANT: Progenics Pharmaceuticals, Inc.

; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED

; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.24

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,460B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/927,931

; FILING DATE: 07-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550

; TELEFAX: (212) 977-9809

; TELEX: 422523 COOP UI

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 530 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: CDNA

; ORIGINAL SOURCE:

; ORGANISM: homo sapien

; CELL TYPE: lymphocyte

US-08-477-460B-4

Query Match 25.3%; Score 469.5; DB 3; Length 530;

Best Local Similarity 32.8%; Pred. No. 2.6e-38;

Matches 121; Conservative 68; Mismatches 129; Indels 51; Gaps 14;

Qy 10 LSLPESGPVTI-----IPPTVKLFHSSCDP-----RGDAHSTIQLCLV 48

Db 174 LELODSGTWCTVQLNQKQVFEKIDIVLAFASFGPSVFPPLAPCSRSTSESTAALGCLV 233

Qy 49 SGFSPAKVHVTV-----LVDGQEAENLFPYTRPKREGGQTFSLQSEVNITQGMSSNTY 104

Db 234 KDYPPEPTVSNWNSGALTSG---VHTFPAVL-----QSSGLYSLSVTVVPSNPF-GTQTY 285

Qy 105 TCHVKH---NGSIFEDSSRRCSDDP-----RGVITYLIPPSPLD-LYENGTPKLTCL 153

Db 286 TCNVDHKPSNTKVDKTVVERKCCVPCPPAPVAGPSVFLFPKPKDTLMISRTPEVTCV 345

Qy 154 VLDLESEE-NITVTWVRKKSIGSASQSKTHHATTSITSLPVDKAWIEGEGYQCR 212

Db 346 VDVSHEDPEVQFNWYDVGVEVHNAKTKPREEQNSTFRVSVLTVVHQDMLNGKEYCK 405

Qy 213 VDHHPFKPIVRSITKLPKRLAPEVYMLPPSPPEE-TGTTTRVTCLIRGFYPSISVQWL 271

Db 406 VSNKGLPAPIEKTISKTKGQPREQVTVLPSPREMTKNQVSLTCLVKGFYPSIAVWE 465

Qy 272 PNNEEDHTGHTTTRPKQDHGTDPSPFLYSRMLVNSKIWEKGNLVTCRVVHEALPGSRTL 331

Db 466 SNGQPEN--NYKTTTPPMLD--SDGSFFLYSKLTVDKSRWQGNVFSVMEALHNNHT- 520

Qy 332 EKSLHYSAG 340

Db 521 QKSLSPG 529

Search completed: July 9, 2003, 13:40:23

Job time : 17 secs

Db 406 VSNKGLPAPIEKTISKTKGQPREQVTVLPSPREMTKNQVSLTCLVKGFYPSIAVWE 465

Qy 272 PNNEEDHTGHTTTRPKQDHGTDPSPFLYSRMLVNSKIWEKGNLVTCRVVHEALPGSRTL 331

Db 466 SNGQPEN--NYKTTTPPMLD--SDGSFFLYSKLTVDKSRWQGNVFSVMEALHNNHT- 520

Qy 332 EKSLHYSAG 340

Db 521 QKSLSPG 529

RESULT 15

US-08-379-516-4

; Sequence 4, Application US/08379516

; Patent No. 6083478

; GENERAL INFORMATION:

; APPLICANT: Allaway, Graham P.

; APPLICANT: Maddon, Paul J.

; TITLE OF INVENTION: Immunoconjugates and Uses Thereof

; TITLE OF INVENTION: Moietty-Conjugated CD4-Gamma2 and CD4-IGG2

; FILE REFERENCE: 41215-A-PCT-US

; CURRENT APPLICATION NUMBER: US/08/379,516

; CURRENT FILING DATE: 1996-06-10

; EARLIER APPLICATION NUMBER: PCT/US93/07422

; EARLIER FILING DATE: 1993-08-06

; EARLIER APPLICATION NUMBER: 07/927,931

; EARLIER FILING DATE: 1992-08-07

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 4

; LENGTH: 530

; TYPE: PRT

; ORGANISM: Homo sapiens

US-08-379-516-4

Query Match 25.3%; Score 469.5; DB 3; Length 530;

Best Local Similarity 32.8%; Pred. No. 2.6e-38;

Matches 121; Conservative 68; Mismatches 129; Indels 51; Gaps 14;

Qy 10 LSLPESGPVTI-----IPPTVKLFHSSCDP-----RGDAHSTIQLCLV 48

Db 174 LELODSGTWCTVQLNQKQVFEKIDIVLAFASFGPSVFPPLAPCSRSTSESTAALGCLV 233

Qy 49 SGFSPAKVHVTV-----LVDGQEAENLFPYTRPKREGGQTFSLQSEVNITQGMSSNTY 104

Db 234 KDYPPEPTVSNWNSGALTSG---VHTFPAVL-----QSSGLYSLSVTVVPSNPF-GTQTY 285

Qy 105 TCHVKH---NGSIFEDSSRRCSDDP-----RGVITYLIPPSPLD-LYENGTPKLTCL 153

Db 286 TCNVDHKPSNTKVDKTVVERKCCVPCPPAPVAGPSVFLFPKPKDTLMISRTPEVTCV 345

Qy 154 VLDLESEE-NITVTWVRKKSIGSASQSKTHHATTSITSLPVDKAWIEGEGYQCR 212

Db 346 VDVSHEDPEVQFNWYDVGVEVHNAKTKPREEQNSTFRVSVLTVVHQDMLNGKEYCK 405

Qy 213 VDHHPFKPIVRSITKLPKRLAPEVYMLPPSPPEE-TGTTTRVTCLIRGFYPSISVQWL 271

Db 406 VSNKGLPAPIEKTISKTKGQPREQVTVLPSPREMTKNQVSLTCLVKGFYPSIAVWE 465

Qy 272 PNNEEDHTGHTTTRPKQDHGTDPSPFLYSRMLVNSKIWEKGNLVTCRVVHEALPGSRTL 331

Db 466 SNGQPEN--NYKTTTPPMLD--SDGSFFLYSKLTVDKSRWQGNVFSVMEALHNNHT- 520

Qy 332 EKSLHYSAG 340

Db 521 QKSLSPG 529

Search completed: July 9, 2003, 13:40:23

Job time : 17 secs